

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 28, 2004, 14:55:01 ; Search time 207 Seconds

(without alignments)
4942.493 Million cell updates/sec

Title: US-09-975-253-1

Perfect score: 2699

Sequence: 1 ggttcagctgccgcagcgc.....aaaaaaaaaaaaaaaaaaaaa 1426

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 4004546

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=framer_n2p.model -DEV=xl
-O=/cgn2_1/USPTO_spool_p/US09975253/runat_28102004_111509_19617/app_query_fasta_1.1607
-DB=A_Geneseq_23Sep04 -QFMT=fastan -SUFFIX=n2p.rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09975253 @CGN_1_1_227 @runat_28102004_111509_19617 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq_23Sep04:*

1: Geneseqp1980s:*

2: Geneseqp1990s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003bs:*

8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2326	86.2	427	3 AAY87783	Aay87783 Human IRF
2	2326	86.2	427	4 AAM78986	Aam78986 Human pro
3	2326	86.2	427	6 ABG76085	Abg76085 Human int
4	2326	86.2	427	7 ABW02028	Abw02028 Human int
5	2326	86.2	427	7 ADG47651	Adg47651 Human IRF
6	2304	85.4	427	2 AAY15102	Aay15102 Modified
7	2299	85.2	455	7 ADG77016	Adg77016 Human nuc
8	2018	74.8	392	4 ABB11963	Abb11963 Human IFN
9	2018	74.8	392	4 AAM79970	Aam79970 Human pro
10	1766.5	65.5	542	2 AAY15104	Aay15104 Chimeric

11	546	20.2	102	3 AAG00953	Aag00953 Human sec
12	504.5	18.7	488	7 ADD01173	Add01173 Human nuc
13	456.5	16.9	516	8 ADJ75465	Adj75465 Marker ge
14	456.5	16.9	516	8 ADJ75413	Adj75413 Marker ge
15	456.5	16.9	516	8 ADL83070	Adl83070 Human PRO
16	456.5	16.9	516	8 ADN04327	Adn04327 Antipsori
17	456	16.9	844	8 ADP29911	Adp29911 Human sec
18	454	16.8	503	2 AAY15103	Aay15103 Modified
19	448	16.6	503	7 ADES8665	Ades8665 Human PRO
20	448	16.6	503	7 ADES8669	Ades8669 Human PRO
21	448	16.6	503	8 ADJ75462	Adj75462 Marker ge
22	448	16.6	503	8 ADJ75410	Adj75410 Marker ge
23	446	16.5	503	4 AAE09329	Aae09329 Human int
24	444.5	16.5	474	8 ADJ75411	Adj75411 Marker ge
25	444.5	16.5	474	8 ADJ75463	Adj75463 Marker ge
26	436	16.2	457	7 ADES8667	Ades8667 Rat Prote
27	436	16.2	457	7 ADES8667	Ades8667 Rat Prote
28	436	16.2	457	8 ADJ76202	Adj76202 Marker ge
29	436	16.2	457	8 ADJ76237	Adj76237 Marker ge
30	399.5	14.8	94	5 ABB57389	Abb57389 Rat mucoc
31	399	14.8	77	4 AAM24435	Aam24435 Human EST
32	397	14.7	450	2 AAR99427	Aar99427 Human Lym
33	397	14.7	450	6 ABU03802	Abu03802 Human exp
34	397	14.7	450	6 ABU03798	Abu03798 Human exp
35	397	14.7	450	8 ADQ18306	Adq18306 Human sof
36	390.5	14.5	451	6 ABU03805	Abu03805 Human exp
37	390.5	14.5	451	6 ABU03799	Abu03799 Human exp
38	390.5	14.5	451	8 ADM67227	Adm67227 Human adi
39	389.5	14.4	449	8 ADM67226	Adm67226 Murine ad
40	389.5	14.4	451	2 AAM38426	Aaw38426 Human mul
41	389.5	14.4	451	6 ABU03804	Abu03804 Human exp
42	389.5	14.4	451	6 ABU03800	Abu03800 Human exp
43	389.5	14.4	451	6 ABU03807	Abu03807 Human exp
44	389.5	14.4	451	6 ABU03803	Abu03803 Human exp
45	389.5	14.4	451	8 ADL83104	Adl83104 Human PRO

ALIGNMENTS

RESULT 1

AAY87783

ID AAY87783 standard; protein; 427 AA.

XX

AC AAY87783;

XX 24-AUG-2000 (first entry)

DT Human IRF3 protein.

DE Human; ADA2; cytostatic; gene therapy; treatment; cancer; IRF3.

XX Homo sapiens.

XX Key Location/Qualifiers

PH Misc-difference 8 /label= unidentified

FT US6054289-A.

XX 25-APR-2000.

XX 30-AUG-1996; 96US-00705771.

XX 30-AUG-1995; 95US-0002993P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Moore PA;

XX WPI; 2000-338491/29.

XX N-PSDB; AAA39474.

XX New polynucleotide encoding human AD2 is useful for treating cancer and

for isolating cDNAs and genes having similar biological activity.

Disclosure; Col 59-62; 54pp; English.

This invention describes a novel polynucleotide (I) encoding human ADA2. The products of the invention have cytosstatic activity and can be used for gene therapy. (I) is useful for treating cancer; as primers and probes for isolating full length cDNA and genes having similar biological activity. This sequence represents the human IRF3 protein described in the method of the invention.

Sequence 427 AA;

ment Scores:	1.9e-166	Length:	427
No.:	2326.00	Matches:	427
Similarity:	100.0%	Conservative:	0
Local Similarity:	100.0%	Mismatches:	0
Match:	68.18%	Indels:	0
	3	Gaps:	0

9-975-253-1 (1-1426) x AAY87783 (1-427)

47 ATGGGAACCCAAAGCAGCGNCTGTCGCCCTGGCTGTGCAGCTGGACTTGGGGCAA 106
1 Met.GlyThr.ProIysProArg***Leu.ProTrp.Leu.Val.Ser.Gln.Leu.Asp.Leu.Gly.Gln 20
107 CTGAGGGCGTGCCCTGGGTGAACAAGACGCGACGGCTTCGCCATCCCTTTGGAGCAC 166
21 Leu.Glu.Gly.Val Ala.Trp.Val.Asn.Lys.Ser.Arg.Thr.Arg.Phe.Arg.Ile.Pro.Trp.Lys.His 40
167 GGCTCATCGCAGGATGCACAGCAGAGGATTTCGGAAATCTTCAGCGCTGGGCCGACGGCC 226
41 Gly.Leu.Arg.Gln.Asp.Alac.Ile.Gln.Glu.Asp.Phe.Gly.Ile.Phe.Gln.Ala.Trp.Pala.Glu.Ala 60
227 ACTGGTGTCATGTTCCCGGGAGGATAGCAGACACTGCCAACCCTGGAGAGGAAATTC 286
61 Thr.Gly.Ala.Trp.Val.Pro.Gly.Arg.Asp.Lys.Pro.Asp.Leu.Pro.Trp.Lys.Arg.Asn.Phe 80
287 CGCTCTGCCCTCAACCGCAAGAAGGGTTGGTTTTAGCAGAGACCGGAGCAGACGACCT 346
81 Arg.Ser.Ala.Leu.Asn.Arg.Lys.Glu.Gly.Leu.Arg.Leu.Ala.Glu.Asp.Arg.Ser.Lys.Asp.Pro 100
347 CAGGACCCACATAAATCTACAGTTTGTAAGCTCAGAGTTCGGGCACTTTTCCCAGCCA 406
101 His.Asp.Pro.His.Lys.Ile.Tyr.Glu.Phe.Val.Asn.Ser.Gly.Val.Gly.Asp.Phe.Ser.Gln.Pro 120
407 GACACCTCTCCGAGACCAATAGTGGAGGAGTACTCTCATACCAGGACAGACACTTCG 466
121 Asp.Thr.Ser.Pro.Asp.Thr.Asn.Gly.Gly.Lys.Trp.Ser.Asp.Trp.Gln.Glu.Asp.Ile.Leu 140
467 GATGAGTTACTGGGTAAATGGTTGTCGCCCCACTCCAGATCCGGGACCCCAAGCCCTG 526
141 Asp.Glu.Leu.Leu.Gly.Asn.Met.Val.Leu.Ala.Pro.Leu.Pro.Asp.Pro.Gly.Pro.Ser.Leu 160
527 GCTGTAGCCCTGAGCCCTGCCTCAGCCCTCGGAGCCCCAGCTTGACAACTCCCACT 586
161 Ala.Val.Ala.Pro.Glu.Pro.Cys.Pro.Gln.Pro.Leu.Arg.Ser.Pro.Ser.Leu.Asp.Asn.Pro.Trp 180
587 CCCTTCCCAACCTGGGGCCCTCTGAGAACCCCACTGAACGGCTGTGTGGTCGGCGGGGAA 646
181 Pro.Phe.Pro.Asn.Leu.Gly.Pro.Ser.Glu.Asn.Pro.Leu.Lys.Arg.Leu.Leu.Val.Pro.Gly.Glu 200
647 GAGTGGGAGTTCCAGGTGCACAGCTTCTACCGGGGGCCCAAGTCTTCAGCAGACCATC 706
201 Glu.Trp.Glu.Phe.Glu.Val.Trp.Ala.Phe.Tyr.Arg.Gly.Arg.Gln.Val.Phe.Gln.Gln.Trp.Ile 220
707 TCTCTCCCGAGGGCGCTCGCGCTGTGGGGTCCGAAGTCGGAGACGAGCACCTGCCTGGA 766
221 Ser.Cys.Pro.Glu.Gly.Leu.Arg.Leu.Val.Gly.Ser.Glu.Val.Gly.Yas.Par.Trp.Leu.Pro.Gly 240
767 TGGCCAGTCACTGCACAGACCTCGGCATGTTCCTGCACAGACGGGGAGTGATGAGTAC 826

Db	241	TtpProValThrLeuProAspProGlyMetSerLeuThrAspArgGlyValMetSerTyr	260
QY	827	GTGAGCGATGTCTCAGCTGCCTGGTGGGGAGCTGGCTCTCTGGCGGCGGGCAGCTGG	886
Db	261	ValArgHisValLeuSerCysLeuGlyGlyLeuAlaLeuTrpArgAlaGlyGlnTrp	280
QY	887	CTCTGGGCCCCAGCGCTGGGGCACTGCCACACACTACTGGCGAGTGAGCGAGGAGCTGCTC	946
Db	281	LeuTrpAlaGlnArgLeuGlyHisCysHisThrTyrTrpAlaValSerGluGluLeuLeu	300
QY	947	CCCAACAGGGCGCATGGCGCTGATGGCGAGTGCCCAAGGACAAGGAAGGAGCGCTGTTT	1006
Db	301	ProAsnSerGlyHisGlyProAspGlyGluValProLysAspLysGluGlyValPhe	320
QY	1007	GACCTGGGGCCCTTCATGTAGACTGATTACCTTCACGGNAGGAAGCGGACGCTCACCA	1066
Db	321	AspLeuGlyProPheIleValAspLeuIleThrPheThrGluGlySerGlyArgSerPro	340
QY	1067	CGCTATGCCCTCTGCTTCTGTGTGGGGAGTCATGGGCCCCAGGACCAAGCGCTGGACCAAG	1126
Db	341	ArgTyrAlaLeuTrpPheCysValGlyGluSerTrpProGlnAspGlnProTrpThrLys	360
QY	1127	AGGCTCGTAGTGTCAGAGTTGTGGCCACAGTGCCTCAGGCGCTTGGTAGAAATGGCCCGG	1186
Db	361	ArgLeuValMetValLysValValProThrCysLeuArgAlaLeuValGluMetAlaArg	380
QY	1187	GTAGGGGTGCCTCCTCCCTGGGAATACTGTGGACCTGCACATTTTCCACAGGCCACCA	1246
Db	381	ValGlyGlyAlaSerSerLeuGluAsnThrValAspLeuHisIleSerAsnSerHisPro	400
QY	1247	CTCTCCCTCACCTCCGACCAGTACAAGGCCCTACTGCAGGACTTGGTGGAGGCATGGAT	1306
Db	401	LeuSerLeuThrSerAspGlnTyrLysAlaTyrLeuGlnAspLeuValGluGlyMetAsp	420
QY	1307	TTCAGGGCCCTGGGGAGAGC	1327
Db	421	PheGlnGlyProGlyGluSer	427
RESULT 2			
AAW78986			
ID	AAW78986 standard; protein; 427 AA.		
XX			
AC	AAW78986;		
XX			
DT	06-NOV-2001 (first entry)		
XX			
DE	Human protein SEQ ID NO 1648.		
XX			
KW	Human; cytokine; cell proliferation; cell differentiation; gene therapy;		
KW	vaccine; peptide therapy; stem cell growth factor; haematopoiesis;		
KW	tissue growth factor; immunomodulatory; cancer; leukaemia;		
KW	nervous system disorder; arthritis; inflammation.		
XX			
OS	Homo sapiens.		
XX			
FN	W0200157190-A2.		
XX			
PD	09-AUG-2001.		
XX			
PF	05-FEB-2001; 2001WO-US004098.		
XX			
PR	03-FEB-2000; 2000US-00496914.		
PR	27-APR-2000; 2000US-00560875.		
PR	20-JUN-2000; 2000US-00598075.		
PR	19-JUL-2000; 2000US-00620325.		
PR	01-SEP-2000; 2000US-00654936.		
PR	15-SEP-2000; 2000US-00663561.		
PR	20-OCT-2000; 2000US-00693325.		
PR	30-NOV-2000; 2000US-00728422.		
XX			
PA	(HYSE-) HYSEQ INC.		
XX			
PI	Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cacy.		

PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Yang Y, Wehrman T, Goodrich R;
 XX
 DR WPI: 2001-476283/51.
 DR N-PSDB; AAKS2119.
 XX
 PT Nucleic acids encoding polypeptides with cytokine-like activities, useful
 PT in diagnosis and gene therapy.
 XX
 PS Claim 20; Page 3986; 6221pp; English.
 XX
 CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAK78323-AAK80302) that exhibit activity relating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
 CC (AAK52582) and 3666 (AAK80020) are omitted as the relevant pages from the
 CC sequence listing were missing at the time of publication
 XX
 SQ Sequence 427 AA;

Alignment Scores:
 Pred. No.: 1.9e-166 Length: 427
 Score: 2326.00 Matches: 426
 Percent Similarity: 99.77% Conservative: 0
 Best Local Similarity: 99.77% Mismatches: 1
 Query Match: 96.18% Indels: 0
 DB: 4 Gaps: 0

US-09-975-253-1 (1-1426) x AAM78986 (1-427)

QY 47 ATGGGAACCCCAAGCAGCGNCTCGCTCGCTGGTGTGCGAGTGCAGCTGGGGCAA 106
 Db 1 MetGlyThrProLysProArgIleLeuProTrpLeuValSerGlnLeuAspLeuGln 20
 QY 107 CTGGAGGGCGTGGCTGGTGAACAGAGCCGCGCTTCGGCATCCCTTGGAGAGCAC 166
 Db 21 LeuGluGlyValAlaTrpValAsnLysSerArgThrArgPheArgIleProTrpLysHis 40
 QY 167 GGCCTACGCGAGATCGACAGCAGGAGGATTTCGGAATCTTCAGCGCTGGGCGGAGGCC 226
 Db 41 GlyLeuArgGlnAspAlaGlnGlnGluAspPheGlyIlePheGlnAlaTrpAlaGluAla 60
 QY 227 ACTGTTGCATATCTCCCGGAGGGATAAGCCAGACCTCGCAACCTGGAGAGGAATTTC 286
 Db 61 ThrGlyAlaTrpValProGlyArgAspLysProAspLeuProThrTrpLysArgAsnPhe 80
 QY 287 CGCTCTCCCTCAACCGCAAGAGGGTTCCGTTTACGAGACCGGACGACGACCCCT 346
 Db 81 ArgSerAlaLeuAsnArgLysGluGlyLeuArgLeuAlaGluAspArgSerLysAspPro 100
 QY 347 CAGCAGCCACATAAATCTACGAGTTGTCACTCAGGAGTTGGGACTTTCGCCAGCCA 406
 Db 101 HisAspProHisLysIleTrpGluPheValAsnSerGlyValGlyAspPheSerGlnPro 120
 QY 407 GACACCTCTCCGACACCAATGTGGAGGAGTACTTCTGATACCCAGGAAGATCTG 466
 Db 121 AspThrSerProAspThrAsnGlyGlySerThrSerAspThrGlnGluAspIleLeu 140
 QY 467 GATGAGTTACTGGTAACTAGTGTTCGCCCACTCCAGATCCGGAGACCCCGAGCTG 526
 Db 141 AspGluLeuLeuGlyAsnMetValLeuAlaProLeuProAspProGlyProProSerLeu 160
 QY 527 GCTGTAGCCCTGAGCCCTGCCCTCAGCCCTCGGAGCCCGCCAGCTTGGACATCCCACT 586
 Db 161 AlaValAlaProGluProCysProGlnProLeuArgSerProSerLeuAspAsnProThr 180

QY 587 CCCTTCCCAACCTGGGGCCCTCTGAGAACCCACTGAAGCGGCTGTGTGGCCGGGAA 646
 Db 181 ProPheProAsnLeuGlyProSerGluAsnProLeuLysArgLeuLeuValProGlyGlu 200
 QY 647 GAGTGGGAGTTCCAGAGTGACAGCTTCTACCGGGGGCGCCAAAGTCTTCCAGCAGACCATC 706
 Db 201 GluTrpGluPheGluValThrAlaPheTrpArgGlyArgGlnValPheGlnGlnThrIle 220
 QY 707 TCCTGCCCGAGGGCTGGCGCTGGTGGCTCCGAGTGGGAGACAGAGCGCTCCCTGGA 766
 Db 221 SerCysProGluGlyLeuArgLeuValGlySerGluValGlyAspArgThrLeuProGly 240
 QY 767 TGGSCAGTCACTGCCAGACCCCTGGCATCTCCCTGACACAGCGGAGTGATGAGTAC 826
 Db 241 TrpProValThrLeuProAspProGlyMetSerLeuThrAspArgGlyValMetSerTyr 260
 QY 827 GTGAGGATGTGTAGCTGCTGGTGGGAGTGTCTCTGGCGGGCGGGCAGCTGG 886
 Db 261 ValArgHisValLeuSerCysLeuGlyGlyLeuAlaLeuTrpArgAlaGlyGlnTrp 280
 QY 887 CTCTGGGGCCCGAGCGCTGGGCGCTGCCACATCTGGGCGAGTGAGCGAGAGTGGCTC 946
 Db 281 LeuTrpAlaGlnArgLeuGlyHisCysHisThrTrpAlaValSerGluGluLeu 300
 QY 947 CCCAACAGCGGGCATGGGCTGTGAGGAGTCCCGAGTCCCGAGGACAGGAGGCGTGT 1006
 Db 301 ProAsnSerGlyHisGlyProAspGlyGluValProLysAspLysGluGlyValPhe 320
 QY 1007 GACCTGGGGCCCTTCATTTGATCTGATTACCTTCAGGAAGGAGGAGCGCTCACCA 1066
 Db 321 AspLeuGlyProPheIleValAspLeuIleThrPheThrGluGlySerGlyArgSerPro 340
 QY 1067 CGCTATGCGCTCTGTTCTGTGGGAGTGTGTCGGGCGGAGTGTGTCGGGCGGAGCCAG 1126
 Db 341 ArgTrpAlaLeuTrpPheCysValGlyGluSerTrpProGlnAspGlnProTrpThrLys 360
 QY 1127 AGGCTCGTGATGCTCAAGGTTGTGCGCCACGCTGCCTCAGGCGCTTGGTGAATGGCCCG 1186
 Db 361 ArgLeuValMetValLysValValProThrCysLeuArgAlaLeuValGluMetAlaArg 380
 QY 1187 GTAGGGGTGGCTCTCTCCCTGGAGAAATCTGTGGACCTGCACATTTCCACAGCCACCA 1246
 Db 381 ValGlyGlyAlaSerSerLeuGluAsnThrValAspLeuHisIleSerAsnSerHisPro 400
 QY 1247 CTCTCCCTCACCTCCGACCGAGTACAGGCTACCTGCGAGGCTTGGTGGAGGCGAT 1306
 Db 401 LeuSerLeuThrSerAspGlnTrpLysAlaTrpLeuGlnAspLeuValGluGlyMetAsp 420
 QY 1307 TTCCAGGGCGCTGGGGAGAGC 1327
 Db 421 PheGlnGlyProGlyGluSer 427
 XX
 RESULT 3
 ABG76085
 ID ABG76085 standard; protein; 427 AA.
 XX
 AC ABG76085;
 XX
 DT 09-MAY-2003 (first entry)
 XX
 DE Human interferon regulatory factor 3, IRF3.
 XX
 KW Human; interferon regulatory factor 3; IRF3; gene therapy; HIV infection;
 KW transcription protein; viral infection; Streptococcus pyogenes;
 KW human T-cell lymphotropic virus type 1; Epstein-Barr virus; scabies;
 KW encephalitis; yellow fever; bacterial infection; Neisseria gonorrhoea;
 KW gingivitis; syphilis; parasitic infection; Amebiasis; Babesiosis;
 KW malaria; deficient serum immunoglobulin production; recurrent infection;
 KW immune system dysfunction; joint infection; bone infection; malignancy;
 KW skin infection; parotid gland infection; blood-borne infection; epilepsy;
 KW autoimmune disease; inflammatory disorder; primary immunodeficiency;
 KW vaccine adjuvant.

```

OS Homo sapiens.
XX Key Location/Qualifiers
XX PF /label= Val, Ile, Leu, Phe
XX FT Misc-difference 8
XX FT /note= "Encoded by NTC"
XX FT Region 141..147
XX FT /label= Nuclear_export_signal
XX FT Domain 198..381
XX FT /label= Interferon_regulatory_factor_association_domain
XX FT Region 382..407
XX FT /label= Phosphorylation_region
XX FT Domain 408..427
XX FT /label= Autoinhibitory_domain
XX XX
XX US2002164694-A1.
XX PN
XX XX
XX PD 07-NOV-2002.
XX XX
XX 12-OCT-2001; 2001US-00975253.
XX XX
XX 30-AUG-1995; 95US-0002993P.
XX PR 30-AUG-1996; 96US-00705771.
XX PR 13-OCT-2000; 2000US-0239936P.
XX XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX XX
XX Moore PA, Pith-Rowe P;
XX PI
XX XX
XX DR WPI; 2003-298697/29.
XX DR N-PSDB; ABX12065.
XX XX
XX New polynucleotide, useful for preventing, treating or ameliorating HIV
XX PT disease.
XX PT
XX Claim 1; Fig 1; 82pp; English.
XX PS
XX XX
XX The invention relates to an isolated human interferon regulatory factor 3
XX CC IRF3 polynucleotide. The interferon regulatory factors are transcription
XX CC proteins. The polynucleotide or polypeptide is useful for preventing,
XX CC treating or ameliorating HIV disease. The polynucleotide or polypeptide
XX CC is also useful for treating viral infections e.g. human T-cell
XX CC lymphotropic virus type I and Epstein-Barr virus and related diseases
XX CC e.g. encephalitis and yellow fever; bacterial infections e.g.
XX CC Streptococcus pyogenes and Neisseria gonorrhoea and related diseases e.g.
XX CC gingivitis and syphilis; parasitic infections e.g. Amebiasis and
XX CC Babesiosis and related diseases e.g. scabies and malaria; deficient serum
XX CC immunoglobulin production; recurrent infections; immune system
XX CC dysfunction; infections of the joints, bones, skin and/or parotid glands;
XX CC blood-borne infections; autoimmune diseases; inflammatory disorders;
XX CC malignancies; primary immunodeficiencies; epilepsy. The polypeptide is also
XX CC useful as a vaccine adjuvant. The present sequence represents the amino
XX CC acid sequence of the human interferon regulatory factor 3, IRF3
XX XX
XX Sequence 427 AA;
XX
XX Alignment Scores:
XX Pred. No.: 1,9e-166 Length: 427
XX Score: 2326.00 Matches: 427
XX Percent Similarity: 100.00% Conservative: 0
XX Best local Similarity: 100.00% Mismatches: 0
XX Query Match: 86.18% Indels: 0
XX DB: 6 Gaps: 0
XX
XX US-09-975-253-1 (1-1426) x ABG76085 (1-427)
XX
XX 47 ATGGGAACCCCAAGCCACGNTCTGCCCTGGCTGGTGTCCGAGCTGCAGCTGGGGCAA 106
XX Db 1 MetGlyThrProLysProArg***LeuProTrpLeuValSerGlnLeuAspLeuGlyGln 20
XX
XX 107 CTGGAGGCGCGCTGGTCAACAGAGCCGACCGCTCCGATCCCTTGAAGCAC 166
XX Db 21 LeuGluGlyValAlaTrpValAsnLysSerArgThrArgPheArgLeuProTrpLysHis 40

```


QY 1247 CTCCTCCTACCTCCGACCAAGTCAAGGCTACCTGAGGACTTGGTGAGGCGCATGGAT 1306
 Db 401 LeuSerLeuThrSerAspGlnTyrLysAlaTyrLeuGlnAspLeuValGluGlyMetAsp 420

QY 1307 TTCAGGGCCCTGGGAGAGC 1327
 Db 421 PheGlnGlyProGlyGluSer 427

RESULT 4
 ABW02028
 ID ABW02028 standard; protein; 427 AA.
 XX AC ABW02028;
 XX AC
 DT 12-FEB-2004 (first entry)
 XX DE Human interferon regulatory factor (IRF3).

XX Human; genetic disease; muscular dystrophy; cystic fibrosis; cytostatic;
 KW scientific research; gene therapy; interferon regulatory factor; IRF3.
 XX OS Homo sapiens.

XX PH Key Location/Qualifiers
 FT Misc-difference 8
 FT FT /note= "Encoded by NTC"

XX US6639052-B1.
 XX 28-OCT-2003.

XX 14-OCT-1999; 99US-00417540.
 XX 30-AUG-1995; 95US-0002993P.
 PR 30-AUG-1996; 96US-00705771.

XX (HUMA-) HUMAN GENOME SCI INC.

XX PA Moore PA;

XX WPI; 2003-842790/78.
 DR N-PSDB; AAD63572.

XX New isolated protein and nucleic acid molecules, useful for diagnostic
 PT and therapeutic purposes, e.g. for treating genetic diseases such as
 PT muscular dystrophy or cystic fibrosis.

XX Example 5; Fig 5; Opp; English.

XX The invention relates to isolated new isolated protein and nucleic acid
 CC molecules useful for diagnostic and therapeutic purposes. The invention
 CC is for treating genetic diseases such as muscular dystrophy or cystic
 CC fibrosis, and for in vitro purposes related to scientific research.
 CC synthesis of DNA and manufacture of DNA vectors. The invention is useful
 CC in gene therapy. The present sequence is human interferon regulatory
 CC factor (IRF3)

XX SQ Sequence 427 AA;

Alignment Scores:
 Pred. No.: Length: 427
 Score: 2326.00 Matches: 427
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 86.18% Indels: 0
 DB: 7 Gaps: 0

US-09-975-253-1 (1-1426) x ABW02028 (1-427)

QY 47 ATGGGAACCCCAAGCCAGGNTCCCTGGCTGGTGGTGGAGCTGGACCTGGGGCAA 106
 Db 1 MetGlyThrProLysProArg***LeuProTrpLeuValSerGlnLeuAspLeuGlyGln 20

QY 107 CTGGAGGGCGTGGCTGGTGAACAAGACCGCACGCGCTTCCGCATCCCTTGGAGCAC 166
 Db 21 LeuGluGlyValAlaTrpValAsnLysSerArgThrArgPheArgIleProTyrLysHis 40
 QY 167 GGCCTACGGCAGGATGCACAGCAGGAGATTTCGGAATCTTCCAGGCTGGGCGGAGGCC 226
 Db 41 GlyLeuArgGlnAspAlaGlnGlnGluAspPheGlyIlePheGlnAlaTrpAlaGluAla 60
 QY 227 ACTGGTGCATATGTTCCCGGAGGGATNAGCCAGACTGCTCCAACTCGGAAGAGAAATTTC 286
 Db 61 ThrGlyAlaTyrValProGlyArgAspLysProAspLeuProThrTrpLysArgAspPhe 80
 QY 287 CGCTCTGCCCTCAACCGCAAGAGGTTGCGTTTAGCAGAGGACCGGACCAAGACCT 346
 Db 81 ArgSerAlaLeuAsnArgLysGluGlyLeuArgLeuAlaGluAspArgSerLysAspPro 100
 QY 347 CACGACCCACATAAAATCTACGAGTTGTGAATCTCAGGAGTTGGGACTTTTCCAGCCA 406
 Db 101 HisAspProHisLysIleTyrGluPheValAsnSerGlyValGlyAspPheSerGlnPro 120
 QY 407 GACACCTCTCGGACACCACTGCTGGAGGAGTACTTCTGATCCCAAGGAGACATTCTG 466
 Db 121 AspThrSerProAspThrAsnGlyGlySerThrSerAspThrGlnGluAspIleLeu 140
 QY 467 GATGAGTTACTGGTAACTGTTGCGCCCACTCCAGATCCGGACCCCAAGCCTG 526
 Db 141 AspGluLeuLeuGlyAsnMetValLeuAlaProLeuProAspProGlyProProSerLeu 160
 QY 527 GCTGTAGCCCTGAGCCCTGCCCTCAGCCCTGCGGAGCCCACTTGGACAATCCCACT 586
 Db 161 AlaValAlaProGluProCysProGlnProLeuArgSerProSerLeuAspAsnProThr 180
 QY 587 CCCTTCCCAACCTGGGGCCCTCTGAGAACCCACTGAAGCGCTGTGTGCGCGGGAA 646
 Db 181 ProPheProAsnLeuGlyProSerGluAsnProLeuLysArgLeuLeuValProGlyGlu 200
 QY 647 GAGTGGGAGTTCGAGGTGACAGCCCTTCTACCGGGCGCCCAAGTCTTCCAGCAGACCATC 706
 Db 201 GluTrpGluPheGluValThrAlaPheTyrArgGlyArgGlnValPheGlnGlnThrIle 220
 QY 707 TCCTGCCCGAGGGCCTGCGGCTGGTGGGTCCGAAGTGGGAGACAGACCGCTGCTGGA 766
 Db 221 SerCysProGluGlyLeuArgLeuValGlySerGluValGlyAspArgThrLeuProGly 240
 QY 767 TGGCCAGTCACTGCGCAGACCCCTGGCATGTCCTGACAGACAGGGAGTGTAGTAC 826
 Db 241 TrpProValThrLeuProAspProGlyMetSerLeuThrAspArgGlyValMetSerTyr 260
 QY 827 GTGAGGCATGTGCTGAGCTGCTGGTGGGGGACTGGCTCTCTGCGGGCGCGGAGTGG 886
 Db 261 ValArgHisValLeuSerCysLeuGlyGlyGlyLeuAlaLeuTrpArgAlaGlyGlnTrp 280
 QY 887 CTCTGGGCGCAGCGGCTGGGGCCTGCCACATCTGGGCGAGTGAGGAGGTGCTC 946
 Db 281 LeuTrpAlaGlnArgLeuGlyHisCysHisThrTyrTrpAlaValSerGluLeuLeuLeu 300
 QY 947 CCCAACAGGGGCATGGGCTGATGGCGAGTCCCAAGGACCAAGGAGGAGCGCGTGT 1006
 Db 301 ProAsnSerGlyHisGlyProAspGlyGluValProLysAspLysGluGlyValPhe 320
 QY 1007 GACCTGGGGCCCTTCATTGTAGATCTGATTACTTTCAGGAAGAGGAGCGCTCACCA 1066
 Db 321 AspLeuGlyProPheIleValAspLeuIleThrPheThrGluGlySerGlyArgSerPro 340
 QY 1067 CGCTATGCCCTCTGGTCTGTGTGGGGAGTGCATGGCCCCAGGACAGCCGTGGACCAAG 1126
 Db 341 ArgTyrAlaLeuTrpPheCysValGlyGluSerTrpProGlnAspGlnProTrpThrLys 360
 QY 1127 AGGCTGCTGATGCTCAAGTTGTGCCCGCTCCCTGAGGCTGTGTGTAAGAAATGCCCGG 1186
 Db 361 ArgLeuValMetValLysValValProThrCysLeuArgAlaLeuValGluMetAlaArg 380
 QY 1187 GTAGGGGTGCTCTCCTCCCTGGAGAATACTGTGGAGCTGCACATTTCACCAAGCCACCA 1246

Db 381 ValGlyGlyAlaSerSerLeuGluAsnThrValAspLeuHisIleSerAsnSerHisPro 400
 QY 1247 CTCTCCCTCACCCTCGACAGCTACAGGCTACCTGACGAGCTGTGTGAGGCGCATGGAT 1306
 Db 401 LeuSerLeuThrSerAspGlnTyrLysAlaTyrLeuGlnAspLeuValGluGlyMetAsp 420
 QY 1307 TTCCAGGGCCCTGGGGAGGAC 1327
 Db 421 PheGlnGlyProGlyGluSer 427

RESULT 5

ID ADG47651
 XX ADG47651 standard; protein; 427 AA.
 AC ADG47651;
 XX 11-MAR-2004 (first entry)
 XX Human IRF3.
 DE muscular dystrophy; cystic fibrosis; hypertension; angina pectoris;
 KW myocardial infarction; ulcer; asthma; allergy; psychosis; depression;
 KW migraine; vomiting; benign prostatic hypertrophy; osteoporosis; human.
 XX Homo sapiens.
 XX US2003208043-Al.
 XX 06-NOV-2003.
 XX 04-JUN-2003; 2003US-00453478.
 XX 30-AUG-1995; 95US-0002993P.
 XX 30-AUG-1996; 96US-00705771.
 XX 14-OCT-1999; 99US-00417540.
 XX (HUMA-) HUMAN GENOME SCI INC.

PI Moore PA, Gentz RL, Ji H, Ni J, Hu J;
 XX WPI; 2003-864796/80.
 DR N-PSDB; ADG47640.
 XX New human polypeptides and polynucleotides, useful for diagnosing or
 PT treating genetic diseases such as muscular dystrophy or cystic fibrosis,
 PT hypertension, asthma, depression or osteoporosis.
 XX Claim 1; SEQ ID NO 16; 56bp; English.
 XX The invention relates to an isolated human polypeptide. The polypeptides,
 CC polynucleotides, agonists or antagonist are useful for diagnosing or
 CC treating genetic diseases such as muscular dystrophy or cystic fibrosis,
 CC hypertension, angina pectoris, myocardial infarction, ulcers, asthma,
 CC allergies, psychoses, depression, migraine, vomiting, benign prostatic
 CC hypertrophy or osteoporosis. The polypeptides and polynucleotides are
 CC useful for in vitro purposes related to scientific research, synthesis of
 CC DNA and manufacture of DNA vector. The present sequence represents the
 CC amino acid sequence of human IRF3.
 XX

SQ Sequence 427 AA;
 Alignment Scores:
 Pred. No.: 1.9e-166 Length: 427
 Score: 2326.00 Matches: 427
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 86.18% Indels: 0
 DB: 7 Gaps: 0

US-09-975-253-1 (1-1426) x ADG47651 (1-427)
 QY 47 ATGGGAACCCCAAGCCAGGNTCTGCCCTGGTGTGTCGACGTGGAGCTGGGGCAA 106

Db 1 MetGlyThrProLysProArg***LeuProThrLeuValSerGlnLeuAspLeuGlyGln 20
 QY 107 CTGAGGGCGTGGCTGGGTGAACAAGAGCCGACGCGCTTCCGATCCCTTTGAGACAC 166
 Db 21 LeuGluGlyValAlaTyrValAsnLysSerArgThrArgPheArgIleProThrLysHis 40
 QY 167 GSCCTACGGCAGGATGCACAGCAGAGGATTTCCGAAATCTTCCAGCCTGGGCCGAGGCC 226
 Db 41 GlyLeuArgGlnAspAlaGlnGlnAspPheGlyIlePheGlnAlaTyrAlaGluAla 60
 QY 227 ACTGTCATATGTTCCCGGAGGATAGCCAGACCTGCCAACCTCTGGAAGGGAATTC 286
 Db 61 ThrGlyAlaTyrValProGlyArgAspLysProAspLeuProThrTrpLysArgAsnPhe 80
 QY 287 CGCTCTGCCCTCAACCGCAAGAGGTTGGCTTTAGCAGAGCAGCGGACGAGCACCT 346
 Db 81 ArgSerAlaLeuAsnArgLysGluGlyLeuArgLeuAlaGluAspArgSerLysAspPro 100
 QY 347 CACGACCCACATAAATCTACGAGTTTGTGAACCTCAGGAGTTGGGACTTTTCCAGCCA 406
 Db 101 HisAspProHisLysIleTyrGluPheValAsnSerGlyValGlyAspPheSerGlnPro 120
 QY 407 GACACCTCTCCGGACACCAATGGTGGAGCAGTACTTCTGATACCCAGGAGACATCTCG 466
 Db 121 AspThrSerProAspThrAsnGlyGlyGlySerThrSerAspThrGlnGluAspIleLeu 140
 QY 467 GATGATTACTGGGTAAACATGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 526
 Db 141 AspGluLeuLeuGlyAsnMetValLeuAlaProLeuProAspProGlyProProSerLeu 160
 QY 527 GCTGTAGCCCTGAGCCCTGCCCTCAGCCCTGGCGGAGCCCGAGCTTGGACATCCCAT 586
 Db 161 AlaValAlaProGluProCysProGlnProLeuArgSerProSerLeuAspAsnProThr 180
 QY 587 CCTTTCCCAAACTGGGGCCCTCTGAGAACCCACTGAAAGCGGCTTGTGGTGGCGGGAA 646
 Db 181 ProPheProAsnLeuGlyProSerGluAsnProLeuLysArgLeuLeuValProGlyGlu 200
 QY 647 GAGTGGAGTTGAGGTGACAGCTTCTACCGGGCCCGCAAGTCTTCCAGCAGACCATC 706
 Db 201 GluTrpGluPheGluValThrAlaPheTyrArgGlyArgGlnValPheGlnGlnThrIle 220
 QY 707 TCCTGCCCGAGGGCTGGGCTGGTGGGTCCGAGTGGGAGCAGCAGCAGCTGGCTCGA 766
 Db 221 SerCysProGluGlyLeuArgLeuValGlySerGluValGlyAspArgThrLeuProGly 240
 QY 767 TGGCCAGTCACTGCCAGACCTGGCATGTCCCTGACAGCAGGGGAGTGATGAGCTAC 826
 Db 241 TrpProValThrLeuProAspProGlyMetSerLeuThrAspArgGlyValMetSerTyr 260
 QY 827 GTGAGCGATGCTGAGTGTGGTGGGAGCTGGTCTCTGGCGGGCGGGCAGTGG 886
 Db 261 ValArgHisValLeuSerCysLeuGlyGlyGlyLeuAlaLeuTrpArgAlaGlyGlnTrp 280
 QY 887 CTCTGGGCCCCAGCGCTGGGCTGCCACATCTGGCAGTGGCAGTGGCAGGAGCTGCTC 946
 Db 281 LeuTrpAlaGlnArgLeuGlyHisCysHisThrTyrTrpAlaValSerGluGluLeuLeu 300
 QY 947 CCCAACGCGGGCATGGGCTGTGGCGAGGTCCCGAGGAGTCCCGAGGAGGAGGAGGCTGTTT 1006
 Db 301 ProAsnSerGlyHisGlyProAspGlyGluValProLysAspLysGluGlyValPhe 320
 QY 1007 GACCTGGGGCCCTTCATTGTAGATCTGATACCTTACGGAAGAGCGGAGCTCACCA 1066
 Db 321 AspLeuGlyProPheIleValAspLeuIleThrPheThrGluGlySerGlyArgSerPro 340
 QY 1067 CGCTATGCCCTCTGGTCTGTGGGGAGTGCATGGCGCCCGAGGACACCCCGTGGACCAAG 1126
 Db 341 ArgTyrAlaLeuTrpPheCysValGlyGluSerTrpProGlnAspGlnProThrThrLys 360
 QY 1127 AGGCTCGTGTGTCAGAGTTGTGCCACGTCCTCAGGGCTTGGTGAATAATGCCCGG 1186

Db 361 ArgLeuValMetValLysValValProThrCysLeuArgAlaLeuValGluMetAlaArg 380
 QY 1187 GTAGGGGGTCCCTCCCTGGAGATACCTGCTGACCTGCACATTTCCACAGCCACCCA 1246
 Db 381 ValGlyGlyAlaSerSerLeuGluAsnThrValAspLeuHisSerAsnSerHisPro 400
 QY 1247 CTCCTCCCTCACCTCCGACCACTACAGGCTTACCTGCAGCACTGTGTGGAGGCGCATGGAT 1306
 Db 401 LeuSerLeuThrSerAspGlnTyrLysAlaTyrLeuGlnAspLeuValGluGlyMetAsp 420
 QY 1307 TTCAGGGCCCTGGGAGAGC 1327
 Db 421 PheGlnGlyProGlyGluSer 427
 RESULT 6
 AAY15102
 ID AAY15102 standard; protein; 427 AA.
 AC AAY15102;
 XX
 DT 25-JAN-2000 (first entry)
 XX
 DE Modified Interferon Regulatory Factor-3 protein (IRF-3).
 XX
 KW Interferon regulatory factor; IRF-3; transactivation domain; serine;
 KW threonine; carboxy terminus; aspartic acid; mutant; phosphorylation;
 KW post-translational modification; sendai virus; cancer treatment; herpes;
 KW activator; promoter; PRDI/PRDIII; ISRE regulatory element; stimulation;
 KW DNA binding; transcriptional activity; transcriptional co-activator;
 KW CBP/p300; proteasome mediated degradation; viral infection; influenza;
 KW HIV infection; activate; cytokine gene; target cell.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 PH Key Location/Qualifiers
 FT Misc-difference 396 /note= "Wild type Ser replaced with Asp"
 FT Misc-difference 398 /note= "Wild type Ser replaced with Asp"
 FT Misc-difference 402 /note= "Wild type Ser replaced with Asp"
 FT Misc-difference 404 /note= "Wild type Ser replaced with Asp"
 FT Misc-difference 405 /note= "Wild type Ser replaced with Asp"
 FT Misc-difference 405 /note= "Wild type Ser replaced with Asp"
 XX
 PN WO9951737-A1.
 XX
 PD 14-OCT-1999.
 XX
 PF 07-APR-1999; 99WO-CA000314.
 XX
 PF 07-APR-1998; 98CA-02234588.
 PR (DAVI-) DAVIS-JEWISH GEN HOSPITAL MORTIMER B.
 XX
 PA Hiscott J, Lin R;
 PI
 XX WPI; 1999-620201/53.
 DR N-PSDB; AAX90992.
 XX
 PT Carboxy-terminus modified highly active forms of interferon regulatory
 factor proteins used for the treatment of viral infections.
 XX
 PS Claim 8; Fig 10; 93pp; English.
 XX
 CC The present protein sequence is the interferon regulatory factor-3 (IRF-
 3) 5D protein, that is modified in the transactivation domain. The serine
 and threonine residues in the carboxy terminus are modified post-
 translationally, by phosphorylation, following sendai virus infection.
 CC The modified IRF-3, substituted with aspartic acid, functions as a strong
 activator of promoters containing PRDI/PRDIII or ISRE regulatory

CC elements, stimulation of DNA binding and transcriptional activity,
 CC association of IRF-3 with the transcriptional co-activator CBP/p300 and
 CC proteasome mediated degradation. The modified IRF proteins are used for
 CC the treatment of viral infections like, influenza, herpes or HIV
 CC infection. They may also be used to activate a cytokine gene, in cancer
 CC treatment or to modify a target cell of an organism
 XX
 SQ Sequence 427 AA;
 Alignment Scores: Length: 427
 Pred. No.: 8.6e-165 Matches: 421
 Score: 2304.00
 Percent Similarity: 98.59% Conservative: 0
 Best Local Similarity: 98.59% Mismatches: 6
 Query Match: 85.36% Indels: 0
 DB: 2 Gaps: 0
 US-09-975-253-1 (1-1426) x AAY15102 (1-427)
 QY 47 ATGGGAACCCCAAGACCCACGNTCTGCTGGTGTGTCGACGTGGACCTGGGGCAA 106
 Db 1 MetGlyThrProLysProArgIleLeuProTyrPheValSerGlnLeuAspLeuGln 20
 QY 107 CTGAGGCGCTGGCTGGGTGAACAAGAGCGCGCTTCCGATCCCTTGAAGCAC 166
 Db 21 LeuGluGlyValAlaTyrPheValAsnLysSerArgThrArgPheArgIleProTyrLysHis 40
 QY 167 GGCCTACGGCAGGATGACAGCAGGAGATTCGAAATCTTCCAGGCTGGCGGCGCC 226
 Db 41 GlyLeuArgGlnAspAlaGlnGlnGluAspPheGlyIlePheGlnAlaTyrPheGlnAla 60
 QY 227 ACTGCTGCATATGTTCCCGGAGGAGGATAGGCACACCTGCCCACTGGAGAGAAATTC 286
 Db 61 ThrGlyAlaTyrValProGlyArgAspLysProAspLeuProThrTyrLysArgAsnPhe 80
 QY 287 CGCTCTGCCCTCAACCCGCAAGAGGGTTCGTTTAGCAGAGACCGGAGCAAGACCT 346
 Db 81 ArgSerAlaLeuAsnArgLysGluGlyLeuArgLeuAlaGluAspArgSerLysAspPro 100
 QY 347 CACGACCCACATAAAATCTACGAGTTTGTGAACCTCAGAGTTGGGAGCTTTTCCAGGCA 406
 Db 101 HisAspProHisLysIleTyrGluPheValAsnSerGlyValGlyAspPheSerGlnPro 120
 QY 407 GACACCTCCGGACACCAATGTTGGAGGAGTACTTCTGATACCCAGAGACATTCCTG 466
 Db 121 AspThrSerProAspThrAsnGlyGlyGlySerThrSerAspThrGlnGluAspLeu 140
 QY 467 GATGAGTTACTGGGTAAACATGTTGTTGCCCCCACTCCAGATCCGGGACCCCAAGCCCTG 526
 Db 141 AspGluLeuLeuGlyAsnMetValLeuAlaProLeuProAspProGlyProProSerLeu 160
 QY 527 GCTGTAGCCCTGAGCCCTGCCCTCAGCCCTGCGGAGCCCGGAGCTGGACATCCCACT 586
 Db 161 AlaValAlaProGluProCysProGlnProLeuArgSerProSerLeuAspAsnProThr 180
 QY 587 CCCTTCCCAACCTCGGGCCCTCTCAGAACCCACCTGAAGCGGCTGTGTGTCGGGGGAA 646
 Db 181 ProPheProAsnLeuGlyProSerGluAsnProLeuLysArgLeuLeuValProGlyGlu 200
 QY 647 GAGTGGAGTTTCGAGGTGACAGCCCTTCTACCGGGGCGCCCAAGTCTTCCAGCAGACATC 706
 Db 201 GluTyrGluPheGluValThrAlaPheTyrArgGlyArgGlnValPheGlnGlnThrIle 220
 QY 707 TCTGCGCGGAGGCGCTCGGCTGTTGGGTCCGAGTGGGAGAGGAGCGCTGCTGGA 766
 Db 221 SerCysProGluGlyLeuArgLeuValGlySerGluValGlyAspArgThrLeuProGly 240
 QY 767 TGGCCAGTCACACTGCCAGACCCCTGCATGTCCTGACACAGAGGAGGAGTGTAGCTAC 826
 Db 241 TrpProValThrLeuProAspProGlyMetSerLeuThrAspArgGlyValMetSerTyr 260
 QY 827 GTGAGGCATGTGCTGAGCTGCTGGGTGGGAGCTGGCTCTCTGCGCGGCGCGGAGTGG 886

Db 261 ValArgHisValLeuSerCysLeuGlyGlyGlyLeuAlaLeuTrpArgAlaGlyGlnTrp 280
 QY 887 CTCTGGCCACGGGCTGGGCTGCTCCACACATAGTGGGAGTGGAGGAGCTGCTC 946
 Db 281 LeuTrpAlaGlnArgLeuGlyHisCysHisThrTyrTrpAlaValSerGluGluLeuLeu 300
 QY 947 CCCAACAGCGGGCATGGGCTGATGGCGAGGTCCCAAGGACAAAGGAGGCGGTGTT 1006
 Db 301 ProAsnSerGlyHisGlyProAspGlyGluValProLysAspLysGluGlyValPhe 320
 QY 1007 GACCTGGGCGCTTCATTAGATCTGATTACCTTCACGGAAGGAGCGAGCTCACCA 1066
 Db 321 AspLeuGlyProPheIleValAspLeuIleThrPheThrGluGlySerGlyArgSerPro 340
 QY 1067 CGCTATGGCTCTGGTCTGTGGGGAGTATGCCCCCAGGACCGGCGGTGGACCAAG 1126
 Db 341 ArgTyrAlaLeuTrpPheCysValGlyGluSerTrpProGlnAspGlnProTrpThrLys 360
 QY 1127 AGGCTGCTGATGTCAGAGTGTGCTCCAGCTGCTCAGGCGCTTGTAGAAATGGCCCGG 1186
 Db 361 ArgLeuValMetValLysValValProThrCysLeuArgAlaLeuValGluMetAlaArg 380
 QY 1187 GTAGGGGCTGCTCTCCCTCGAGAAATACTGTGGACCTGCATTTCCAAACGCCACCA 1246
 Db 381 ValGlyGlyAlaSerSerLeuGluAsnThrValAspLeuHisIleAspAsnAspHisPro 400
 QY 1247 CTCTCCCTCACCTCCGACAGTACAGGCTACCTGACGAGCTTGGTGGAGGCGATGGAT 1306
 Db 401 LeuAspLeuAspAspGlnTyrLysAlaTyrLeuGlnAspLeuValGluGlyMetAsp 420
 QY 1307 TTCCAGGGCGCTGGGAGAGC 1327
 Db 421 PheGlnGlyProGlyGluSer 427

RESULT 7

ADG77016
 ID ADG77016 standard; protein; 455 AA.
 AC ADG77016;
 XX
 DT 11-MAR-2004 (first entry)
 XX
 DE Human nucleic acid associated polypeptide (NAAP) 44.
 XX
 KW nucleic acid associated polypeptide; NAAP; cytostatic;
 KW antiarteriosclerotic; anticonvulsant; neurotropic; neuroprotective;
 KW cerebroprotective; anti-HIV; anti-allergic; anti-inflammatory;
 KW thyromimetic; gene therapy; cell proliferative; cancer; atherosclerosis;
 KW neurological; epilepsy; Huntington's disease; stroke; immune;
 KW inflammatory; AIDS; allergy; developmental; Hypothyroidism;
 KW Cushing's syndrome; infection; human.
 XX
 OS Homo sapiens.
 XX
 XX WO2003076586-A2.
 XX
 XX 18-SEP-2003.
 XX
 XX 05-MAR-2003; 2003WO-US007002.
 XX
 XX 06-MAR-2002; 2002US-0362329P.
 XX
 XX 15-MAR-2002; 2002US-0364438P.
 XX
 XX 19-APR-2002; 2002US-0373891P.
 XX
 XX (INCY-) INCYTE GENOMICS INC.
 XX
 XX Richardson TW, Elliott VS, Sprague WW, Jiang X, Tang YT;
 XX Zebajadian Y, Baughn MR, Jackson AA, Kable AB, Lee EA, Khan PA;
 XX Yue H, Forsythe IJ, Ramkumar J, Griffin JA, Li JX, Marquis JP;
 XX Lehr-Mason PM, Jin P, Hawkins PR, Wilson AD, Swarnakar A, He A;
 XX Hafalia AJA, Tran B, Duggan BM, Emerling BM, Borowsky ML, Yao MG;
 XX Chawla NK, Burford N, Khare R, Lee S, Becha SD, Lee SY;

DR WPI: 2003-756814/71.
 DR N-PSDB; ADG77069.
 XX
 PT New human nucleic acid associated proteins (NAAP), useful for diagnosing,
 PT treating and preventing diseases or conditions associated with the
 PT aberrant NAAP expression e.g. cancer, AIDS, atherosclerosis, epilepsy, or
 PT infections.
 XX
 PS Claim 1; SEQ ID NO 44; 373pp; English.
 XX
 CC This invention relates to a novel isolated human nucleic acid associated
 CC polypeptide (NAAP). The invention may be useful for the development of
 CC compounds with a cytostatic, antiarteriosclerotic, anticonvulsant,
 CC neurotropic, neuroprotective, cerebroprotective, anti-HIV, anti-allergic,
 CC anti-inflammatory or thyromimetic activity and also for gene therapy. The
 CC invention may prove useful for the development of treatments for diseases
 CC or conditions associated with the decreased expression or overexpression
 CC of NAAP, such as cell proliferative (for example cancer,
 CC atherosclerosis), neurological (for example epilepsy, Huntington's
 CC disease, stroke), immune/inflammatory (for example AIDS, allergies) and
 CC developmental (for example Hypothyroidism, Cushing's syndrome) disorders,
 CC or infections. The present sequence is that of a human NAAP protein of
 CC the invention.
 XX
 SQ Sequence 455 AA;
 Alignment Scores:
 Pred. No.: 2,08e-164 Length: 455
 Score: 2299.00 Matches: 425
 Percent Similarity: 93.63% Conservative: 1
 Best Local Similarity: 93.41% Mismatches: 1
 Query Match: 85.18% Indels: 28
 DB: 7 Gaps: 1
 US-09-975-253-1 (1-1426) x ADG77016 (1-455)
 QY 47 ATGGAAACCCCAAGCCACGNTCTCGCTGGTGTGTCAGCTGCAGCTGGCGGAA 106
 Db 1 MetGlyThrProLysProArgIleLeuProTrpLeuValSerGlnLeuAspLeuGln 20
 QY 107 CTGGAGGCGTGGCTGGTGAACAAAGCGCGCACCGCTTCCGATCCCTTGGAGCAC 166
 Db 21 LeuGluGlyValAlaTrpValAsnLysSerArgThrArgPheArgIleProTrpLysHis 40
 QY 167 GGCCTACCGCAGGATGCACAGCAGGATTTCCGAATCTCCAGGCTGGCCGAGGCC 226
 Db 41 GlyLeuArgGlnAspAlaGlnGlnGluAspPheGlyIlePheGlnAlaTrpAlaGluAla 60
 QY 227 ACTGTGCATATGTTCCCGGAGGAGTAAAGCAGACCTGCCAACCTGGAAGAGGAATTTC 286
 Db 61 ThrGlyAlaTyrValProGlyArgAspLysProAspLeuProThrTrpLysArgAsnPhe 80
 QY 287 CGCTCTGCCCTCAACCCGCAAGAGGGTTGGTTAGCAGAGGACCGGAGCAAGCACCT 346
 Db 81 ArgSerAlaLeuAsnArgLysGluGlyLeuArgLeuAlaGluAspArgSerLysAspPro 100
 QY 347 CACGACCCACATAAAATCTACGAGTTGTGTAACCTCA----- 382
 Db 101 HisAspProHisLysIleTyrGluPheValAsnSerGlyLeuProAlaTrpGlySerCys 120
 QY 382 ----- 382
 Db 121 GlyArgLeuGlyProGlnGlyPheLeuAlaLeuLeuProSerProLeuPheSerThr 140
 QY 383 GGAGTTGGGAGCTTTCCAGCCAGACACCTCTCCGGACACCAATGGTGGAGGCGAGTACT 442
 Db 141 GlyValGlyAspPheSerGlnProAspThrSerProAspThrAsnGlyGlyGlySerThr 160
 QY 443 TCTGATACCCAGGAGACATTTCTGATGAGTTACTGGGTAAACATGTTGCCCCCCTC 502
 Db 161 SerAspThrGlnGluAspIleLeuAspGlnLeuLeuGlyAsnMetValLeuAlaProLeu 180
 QY 503 CCAGATCCGGACCCCCCAAGCGCTGGCTGTAGCCCTTGAGCCCTGCCTTCAGCCCTCGCG 562

Db 181 ProAspProGlyProProSerLeuAlaValAlaProGluProCysProGlnProLeuArg 200
QY 563 AGCCCGAGCTGGACAAATCCACTCCCTCCCAAACTGGGGCCCTCAGAACCCACTG 622
Db 201 SerProSerLeuAspAsnProThrProPheProAsnLeuGlyProSerGluAsnProLeu 220
QY 623 AGCGGCTGTGGTGGGGGGAGAGTGGAGATTGAGGTGACACGCTTCTACCGGGC 682
Db 221 LysArgLeuValProGlyGluGluTrpGluPheGluValThrAlaPheTyrArgGly 240
QY 683 CGCAAGTCTTCAGCAGACCATCTCTCCCGAGGGCTGGGGTGGGTCGGA 742
Db 241 ArgGlnValPheGlnThrIleSerCysProGluGlyLeuValGlySerGlu 260
QY 743 GTGGGAGACGAGCGCTGGCTGGAGTGGCCAGTCCACACTGCGACCCCTGGCATGTCCTG 802
Db 261 ValGlyAspArgThrLeuProGlyTyrProValThrLeuProAspProGlyMetSerLeu 280
QY 803 ACAGACGGGAGTGTAGCTACGTGAGCATGTCTGAGTGCCTGGTGGGGAGCTG 862
Db 281 ThrAspArgGlyValMetSerTyrValArgHisValLeuSerCysLeuGlyGlyLeu 300
QY 863 GCTCTCTGGCGCGCGGAGTGGCTGGGCCCGAGCGCTGGGGCTGGGCTGGGTCGGA 922
Db 301 AlaLeuTrpArgAlaGlyGlnTrpLeuTrpAlaGlnArgLeuGlyHisCysHisThrTyr 320
QY 923 TGGGAGTGGAGGAGGAGTGTCTCCCAACAGCGGGCATGGGCTGATGGCGAGGTCCCG 982
Db 321 TrpAlaValSerGluGluLeuLeuProAsnSerGlyHisGlyProAspGlyGluValPro 340
QY 983 AAGGACAGCAGAGGAGCGCTGTGACCTGGGGCCCTTATTGTAGATCTGATTACCTTC 1042
Db 341 LysAspLysGluGlyGlyValPheAspLeuGlyProPheIleValAspLeuIleThrPhe 360
QY 1043 ACGGAGGAGGAGCGCTGACACAGCTATGCCCTGCTGCTGCTGGTGGGGAGTCACTG 1102
Db 361 ThrGluGlySerGlyArgSerProArgTyrAlaLeuTrpPheCysValGlyGluSerTrp 380
QY 1103 CCCAGGACCGCGTGGACCAAGAGGCTGCTGATGGTCAAGGTGTGGCCAGTCCCTC 1162
Db 381 ProGlnAspGlnProTrpThrLysArgLeuValMetValLysValValProThrCysLeu 400
QY 1163 AGGGCTTGTAGAAATGCGCGGTAGGGGTGCTCTCTCTGAGAAATCTGTGGAC 1222
Db 401 ArgAlaLeuValGluMetAlaArgValGlyAlaSerSerLeuGluAsnThrValAsp 420
QY 1223 CTGCACATTTCAACAGCAGCCACTCTCCCTCACCTCGACACAGGTACAGGCTACCTG 1282
Db 421 LeuHisIleSerAsnSerHisProLeuSerLeuThrSerAspGlnTrpLysAlaTyrLeu 440
QY 1283 CAGGACTTGGTGGAGGCGATGATTTCCAGGGCCCTGGGGAGAGC 1327
Db 441 GlnAspLeuValGluGlyMetAspPheGlnGlyProGlyGluThr 455
RESULT 8
ID ABB11963
XX ABB11963 standard; peptide; 392 AA.
AC ABB11963;
XX
DT 11-JAN-2002 (first entry)
XX
DE Human IFN regulatory factor homologue, SEQ ID NO:2333.
KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW chronic inflammatory condition; proliferative retinopathy;
KW atherosclerosis; coronary heart disease; arterial ischaemia;
KW bone disorder; osteoporosis; vascular growth disorder;

KW tissue regeneration; wound healing; infection; immune disorder;
KW cell culture; drug screening; gene therapy; antiinflammatory;
KW cytoasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
KW cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KW antifungal; vulnery; antiulcer.
OS Homo sapiens.
XX WO200157188-A2.
PN 09-AUG-2001.
PD 05-FEB-2001; 2001WO-US003800.
XX 03-FEB-2000; 2000US-00496914.
PR 27-APR-2000; 2000US-00560875.
XX (HYSE-) HYSEQ INC.
PA Tang YT, Liu C, Drmanac RT;
PI WPI; 2001-457740/49.
XX N-PSDB; ABA09207.
DR Human proteins and DNA encoding sequences useful for preventing, treating
PT or ameliorating a medical condition in a mammalian subject e.g. arthritis
PT and cancer.
PS Claim 20; Page 287-288; 1963pp; English.
XX Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
CC invention also relates to vectors and recombinant host cells comprising a
CC nucleotide of the invention, methods of producing the novel polypeptides,
CC antibodies against the polypeptides, methods of detecting the nucleotides
CC or polypeptides in a sample, and methods of identifying compounds which
CC bind to polypeptides of the invention. Although novel, many of the
CC polypeptides of the invention have homology to known proteins, and hence
CC giving an insight into their probable biological activities, and hence
CC potential therapeutic applications. The polypeptides of the invention may
CC have various activities, including cytokine, cell proliferation or cell
CC differentiation activities; stem cell growth factor activity;
CC haematopoiesis regulatory activity; tissue growth activity;
CC immunomodulatory activity; activin- or inhibin-related activities;
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC thrombolytic activities; receptor or ligand activities; or may be
CC involved in oncogenesis, cancer cell proliferation or metastasis.
CC Depending on their biological activities, polypeptides and nucleotides of
CC the invention are useful for preventing, treating or ameliorating medical
CC conditions, e.g., by protein or gene therapy. Such conditions include
CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC proliferative retinopathy, atherosclerosis, coronary heart disease,
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
CC vascular growth. Polypeptides involved with tissue regeneration and
CC repair (or nucleic acids encoding them) may be used to promote wound
CC healing (e.g., of burns, incisions and ulcers), while those with
CC immunomodulatory activities may be used in the treatment of viral,
CC bacterial and fungal infections in addition to immune disorders.
CC Polypeptides with growth factor activity may be used in cell cultures to
CC promote cell growth. For example, such polypeptides may be used to
CC manipulate stem cells in culture to give rise to neuroepithelial cells
CC that can be used to augment or replace cells damaged by illness,
CC autoimmune disease or accidental damage. The polypeptides and nucleotides
CC may also be used in the diagnosis of the above conditions, and in drug
CC screening techniques. The present sequence represents a novel human
CC polypeptide of the invention
SQ Sequence 392 AA;
Alignment Scores: 2.91e-143 Length: 392
Pred. No.: 2018.00 Matches: 372
Score:

Percent Similarity:	98.41%	Conservative:	0
Best Local Similarity:	98.41%	Mismatches:	6
Query Match:	74.77%	Indels:	1
DB:	4	Gaps:	1
US-09-975-253-1 (1-1426) x ABB11963 (1-392)			
QY	212	GCTGGGCGGAGCCACCTGTCATATGTTCCCGGAGGGATAAGCCAGACCTGCCAACC	271
Db	15	AlaTprAlaGluAlaThrGlyAlaTyrValProGlyArgAspLysProAspLeuProThr	34
QY	272	TGGAAGAGCAATTTCCGCTCTCCCTCAACCCCAAGAGGGTGTGTTAGCAGAGGAC	331
Db	35	TrpLysArgAsnPheArgSerAlaLeuAsnArgLysGluGlyLeuAlaGluAsp	54
QY	332	CGGAGCAAGACCTCACCACCCACATATAATCTACGAGTTTGTGAACCTCAGAGTTGGG	391
Db	55	ArgSerLysAspProHisAspProHisLysLysLysLysLysLysLysLysLysLys	74
QY	392	GACTTTTCCGAGCCAGACACCTCTCCGGACCAATGGTGGAGGAGTACTTCTGATACC	451
Db	75	AspPheSerGlnProAspThrSerProAspThrAsnGlyGlySerThrSerAspThr	94
QY	452	CAGGAGACATCTTGATGAGTTACTGGGTAAACATGTTGGCCCACTCCCGATCCG	511
Db	95	GlnGluAspLeuAspGluLeuLeuGlyAsnMetValLeuAlaProLeuProAspPro	114
QY	512	GGACCCCAAGCCTGGCTGTAGCCCTCAGCCCTGAGCCCTGAGCCCTGAGCCCTGAG	571
Db	115	GlyProProSerLeuAlaValAlaProGluProCysProGlnProLeuArgSerProSer	134
QY	572	TTGGCAATCCCACTCCCTTCCCAACCTGGGCGCTCTGAGAACCCACTGAAGCGCTG	631
Db	135	LeuAsnProThrProPheProAsnLeuGlyProSerGluAsnProLeuLysArgLeu	154
QY	632	TTGGTCCGGGAGAGTGGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAGTTCGAG	691
Db	155	LeuValProGlyGluGluTrpGluPheGluValThrAlaPheTyrArgGlyArgGluVal	174
QY	692	TTCCAGCAGACCACTCTCTGCGCGGAGGCTGCGGCTGGTGGGTCGAGTGGGAGAC	751
Db	175	PheGlnGlnThrLysSerCysProGluGlyLeuArgLeuValGlySerGluValGlyAsp	194
QY	752	AGACGCTGCTGATGCGGATGTCACACTGCGCAGCCCTGCGATGTCCTGACAGCAGG	811
Db	195	ArgThrLeuProGlyTyrProValThrLeuProAspProGlyMetSerLeuThrAspArg	214
QY	812	GGAGTGATGAGTACGTGAGGAGTGTGCTGAGCTGCTGCTGGTGGGAGTGGCTCTCTGG	871
Db	215	GlyValMetSerTyrValArgHisValLeuSerCysLeuGlyGlyGlyLeuAlaLeuTyr	234
QY	872	CGGCGCGGAGTGGCTCTGGGCGGAGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG	931
Db	235	ArgAlaGlyGlnTyrLeuTyrAlaGlnArgLeuGlyHisCysHisThrTyrTyrAlaVal	254
QY	932	AGCAGAGAGTCTCTCCCAACAGCGGCGATGGGCTGATGCGGAGTCCCGCAGGACAG	991
Db	255	SerGluGluLeuLeuProAsnSerGlyHisGlyProAspGlyGluValProLysAspLys	274
QY	992	GAAGGAGCGTGTGGCTGAGCTGGGCGCTTCTTCTGTA-----GATCTG	1033
Db	275	GluGlyGlyValPheAspLeuGlyProPheLeuValGlySerLeuGlyProProAspLeu	294
QY	1034	ATTACCTTCACGAGGAGGAGCGCTCACAGCTATGCTGCTGCTGCTGCTGCTGCTG	1093
Db	295	IleThrPheThrGluGlySerGlyArgSerProArgTyrAlaLeuTyrPheCysValGly	314
QY	1094	GAGTCATGCGCCAGGACCGCTGGACCAAGAGGCTCGATGCTGCTGCTGCTGCTGCTG	1153
Db	315	GluSerTyrProGlnAspGlnProTyrThrLysArgLeuValMetValLysValPro	334
QY	1154	ACGTGCTTCAGGGCGCTTGGTAGAAATGGCCCGGCTAGGGGTGCTCTCTCTGAGAA	1213

Db	335	ThrCysLeuArgAlaLeuValGluMetAlaArgValGlyGlyAlaSerSerLeuGluAsn	354
QY	1214	ACTGTGGACCTGCACATTTCCAAACAGCCACCTCTCCCTCACCTCCGACGACGACG	1273
Db	355	ThrValAspLeuHisLysSerAsnSerHisProLeuSerLeuThrSerAspGlnTyrLys	374
QY	1274	GCTTACCTGCAGGACTTGTGTGGAGGCGCATGTTTCCAGGGCCCTGGGGAGAGC	1327
Db	375	AlaTyrLeuGlnAspLeuValGluGlyMetAspPheGlnGlyProGlyGluSer	392

RESULT 9

AAM79970

ID AAM79970 standard; protein; 392 AA.

XX

AC AAM79970;

XX

06-NOV-2001 (first entry)

XX

Human protein SEQ ID NO 3616.

DE

Human; cytokine; cell proliferation; cell differentiation; gene therapy; vaccine; peptide therapy; stem cell growth factor; haematopoiesis; tissue growth factor; immunomodulatory; cancer; leukaemia; nervous system disorder; arthritis; inflammation.

OS

Homo sapiens.

XX

WO200157190-A2.

FN

XX

09-AUG-2001.

XX

05-FEB-2001; 2001WO-US004098.

XX

03-FEB-2000; 2000US-00496914.

XX

27-APR-2000; 2000US-00560875.

XX

20-JUN-2000; 2000US-00598075.

XX

19-JUL-2000; 2000US-00620325.

XX

01-SEP-2000; 2000US-00654936.

XX

15-SEP-2000; 2000US-00663561.

XX

20-OCT-2000; 2000US-00693325.

XX

30-NOV-2000; 2000US-00728422.

XX

(HYSE-) HYSEQ INC.

XX

Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;

PI

Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;

PI

Xue AU, Yang Y, Wejhrman T, Goodrich R;

XX

WPI; 2001-476283/51.

XX

N-PSDB; AAKS3103.

XX

Nucleic acids encoding polypeptides with cytokine-like activities, useful in diagnosis and gene therapy.

PT

Claim 20; Page 397; 6221pp; English.

XX

The invention relates to polynucleotides (AAK51456-AAK53435) and the encoded polypeptides (AAM78323-AAM80302) that exhibit activity relating to cytokine, cell proliferation or cell differentiation or which may induce production of other cytokines in other cell populations. The polynucleotides and polypeptides are useful in gene therapy, vaccines or peptide therapy. The polypeptides have various cytokine-like activities, e.g. stem cell growth factor activity, haematopoiesis regulating activity, tissue growth factor activity, immunomodulatory activity and activin/inhibin activity and may be useful in the diagnosis and/or treatment of cancer, leukaemia, nervous system disorders, arthritis and inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the sequence listing were missing at the time of publication

XX

Sequence 392 AA;

SQ

Alignment Scores:

Pred. No.:		2,91e-143	Length:	392
Score:		2018.00	Matches:	372
Percent Similarity:		98.41%	Conservative:	0
Best Local Similarity:		98.41%	Mismatches:	0
Query Match:		74.77%	Indels:	6
DB:		4	Gaps:	1
US-09-975-253-1 (1-1426) x AAM79970 (1-392)				
QY	212	GCCTGGCGGAGCCACTGGTGCATATGTTCCGGGAGGAGTAAAGCCAGACCTGCCAAC	271	
Db	15	AlaTrpAlaGluAlaThrGlyAlaTyrValProGlyArgAspLysProAspLeuProThr	34	
QY	272	TGGAGAGCAATTCGGCTCTCCCTCAACCCCAAGAGGGTGGTTCAGCAGAGAC	331	
Db	35	TrpLysArgAspPheArgSerAlaLeuAsnArgLysGluGlyLeuArgLeuAlaGluAsp	54	
QY	332	CGGAGCAAGGACCTCAGCAGCCACATAAAATCAGCAGTTTGTGAATCAGGAGTTGG	391	
Db	55	ArgSerLysAspProHisAspProHisLysLysLysLysLysLysLysLysLysLys	74	
QY	392	GACTTTTCCAGCAGACCTCTCCGACACCAATGGTGGAGGAGTACTTCTGTATCC	451	
Db	75	AspPheSerGlnProAspThrSerProAspThrAsnGlyGlySerThrSerAspThr	94	
QY	452	CAGGAAGACATCTGGATGAGTACTGGCTAACATGGTGTGGCCCACTCCAGATCCG	511	
Db	95	GlnGluAspLysLeuAspLysLeuGlyAsnMetValLeuAlaProLeuProAspPro	114	
QY	512	GGACCCCAAGCTGGCTGTAGCCCTGAGCCCTGACCCCTGAGCCCTGGGAGCCCCAG	571	
Db	115	GlyProSerLeuAlaValAlaProGluProCysProGlnProLeuArgSerProSer	134	
QY	572	TTGGACATCCACCTCCCTCCCAACCTGGGGCCCTCTGAGAACCCACTGAAGCGCTG	631	
Db	135	LeuAspAsnProThrProPheProAsnLeuGlyProSerGluAsnProLeuLysArgLeu	154	
QY	632	TTGGTGGCGGGGAGAGTGGAGTTCGAGGTGACAGCTTCTACGGGGCCGCAAGTC	691	
Db	155	LeuValProGlyGluGluTrpGluPheGluValThrAlaPheTyrArgGlyArgGlnVal	174	
QY	692	TTCCAGCAGACCATCTCTCCCGGAGGCGCTGGCTGGGTCCGAAGTGGGAGAC	751	
Db	175	PheGlnGlnThrLysSerCysProGluGlyLeuArgLeuValGlySerGluValGlyAsp	194	
QY	752	AGGAGCTGCTGGATGCGCAGTCACATGCGCAGACCTGGCATGTCCTGACAGACAGG	811	
Db	195	ArgThrLeuProGlyTrpProValThrLeuProAspProGlyMetSerLeuThrAspArg	214	
QY	812	GGAGTGATGAGTACGTGAGGCATGCTGAGCTGCCTGGGTGGGGAGTGGCTCTCTGG	871	
Db	215	GlyValMetSerTyrValArgHisValLeuSerCysLeuGlyGlyLeuAlaLeuTrp	234	
QY	872	CGGGCCGGCAGTGGCTCTGGGCCCGGCGCTGGGGCTGACACACATCTGGGCACTG	931	
Db	235	ArgAlaGlyGlnTrpLeuTrpAlaGlnArgLeuGlyHisCysHisThrTyrTrpAlaVal	254	
QY	932	AGCGAGGAGTGTCTCCCAAGCGGGGATGGGCTGTGAGTGGAGGTCCCAAGGACAG	991	
Db	255	SerGluGluLeuLeuProAsnSerGlyHisGlyProAspGlyGluValProLysAspLys	274	
QY	992	GAAGGAGCGTGTGTGACCTGGGGCCCTTCATGTA-----GATCTG	1033	
Db	275	GluGlyGlyValPheAspLeuGlyProPheIleValGlySerLeuGlyProProAspLeu	294	
QY	1034	ATTACCTTCACGAGGAGGAGCGCTCACCACGCTATGCCCTCTGCTCTGTGTGGGG	1093	
Db	295	IleThrPheThrGluGlySerGlyArgSerProArgTyrAlaLeuTrpPheCysValGly	314	
QY	1094	GAGTCATGCCCCAGGACCGCGTGGACCAAGAGGCTCGTGATGGTCAAGGTTGTGCC	1153	
Db	315	GluSerTrpProGlnAspGlnProTrpThrLysArgLeuValMetValLysValPro	334	
QY	1154	ACGTGCTCAGGCGCTTGTGTAGAAATGGCCCGGTAGGGGGTCCCTCTCTCTGGAGAT	1213	
Db	335	ThrCysLeuArgAlaLeuValGluMetAlaArgValGlyGlyAlaSerSerLeuGluAsn	354	
QY	1214	ACTGTGGACCTGCACATTTCCACACGCCACCTCTCTCTCTCTCTCTCTCTCTCTCT	1273	
Db	355	ThrValAspLeuHisSerAsnSerHisProLeuSerLeuThrSerAspGlnTyrLys	374	
QY	1274	GCCTACTCAGGACTTGTGTGGAGGCGATGTTTCCAGGGCCCTGGGAGAGC	1327	
Db	375	AlaTyrLeuGlnAspLeuValGluGlyMetAspPheGlnGlyProGlyGluSer	392	
RESULT 10				
AAV15104				
ID	AAV15104	standard; protein; 542 AA.		
XX	AAV15104;			
XX	25-JAN-2000	(first entry)		
XX	Chimeric protein	IRF-7(1-246)/IRF-3(132-427).		
XX	Interferon regulatory factor; IRF; chimeric protein; serine; threonine;			
XX	carboxy terminus; amino terminal domain; aspartic acid; phosphorylation;			
XX	PRDI/PRDIII; promoter; ISRE regulatory element; stimulation; activate;			
XX	activator; DNA binding; transcriptional activity; viral infection;			
XX	proteasome mediated degradation; influenza; HIV infection; cytokine gene;			
XX	target cell.			
OS	Homo sapiens.			
XX	Synthetic.			
XX	Key	Location/Qualifiers		
FT	Region	1..246		
FT		/note= "Corresponds to modified amino-terminal domain of		
FT		IRF-7 protein"		
FT	Region	247..542		
FT		/note= "Corresponds to modified carboxy-terminus of IRF-3		
FT		protein"		
FT	Misc-difference	511		
FT		/note= "Wild type Ser replaced with Asp"		
FT	Misc-difference	513		
FT		/note= "Wild type Ser replaced with Asp"		
FT	Misc-difference	517		
FT		/note= "Wild type Ser replaced with Asp"		
FT	Misc-difference	519		
FT		/note= "Wild type Ser replaced with Asp"		
FT	Misc-difference	520		
FT		/note= "Wild type Ser replaced with Asp"		
XX	WO9951737-A1.			
XX	14-OCT-1999.			
XX	07-APR-1999;	99WO-CA000314.		
XX	07-APR-1998;	98CA-02234588.		
XX	(DAVI-) DAVIS-JEWISH GEN HOSPITAL MORTIMER B.			
XX	Hiscott J, Lin R;			
XX	WPI; 1999-620201/53.			
XX	N-PSDB; AAX90994.			
XX	Carboxy-terminus modified highly active forms of interferon regulatory			
XX	factor proteins used for the treatment of viral infections.			
XX	Claim 11; Fig 13; 93pp; English.			
XX	The present sequence is the chimeric protein comprising, residues 1-246			
XX	from the modified amino-terminal domain of interferon regulatory factor,			

IRF-7 and 132-427 residues from the carboxy-terminus of modified IRF-3 (5D) protein. The serine and threonine residues are post-translationally modified by phosphorylation, following sendai virus infection. The modified IRF, substituted with aspartic acid, functions as a strong activator of promoters containing ISRE and PRDI/PRDIII regulatory elements, stimulation of DNA binding and transcriptional activity and proteasome mediated degradation. The modified IRF sequences are used for the treatment of viral infections like, influenza, herpes or HIV infection. They may also be used to activate a cytokine gene, in cancer treatment or to modify a target cell of an organism

Sequence 542 AA;

Alignment Scores:	
Pred. No.:	2,796-124
Score:	1766.50
Percent Similarity:	67.73%
Best Local Similarity:	64.54%
Query Match:	65.45%
bp	2
Length:	542
Matches:	344
Conservative:	17
Mismatches:	61
Indels:	111
Gaps:	6

US-09-075-253-1 (1-1426) x AAY15104 (1-542)

QY	62	CCACGGNTCTTG-----CCCTGGCTGGTGTGCGAGCTGGACCTGGGGCAACTGGAGGGC	115
DB	10	ProArgValLeuPheGlyGluTrpLeuLeuGlyGluIleSerSerGlyCysTyArgGluGly	29
QY	116	GTGGCTGGGTGAACAAGAGCGGCACGGCTTCGGCATCCCTTGGAGCACGGCTACGG	175
DB	30	LeuGlnTrpLeuAspGluAlaArgThrCysPheArgValProTrpLysHisPheAlaArg	49
QY	176	CAGGATGCACAGCAGGAGATTTCGGAATCTTCCAGGCTGGGCCGAGCCCACTGGTCA	235
DB	50	LysAspLeuSerGluAlaAspAlaArgIlePheLysAlaTrpAlaValAlaArgGlyArg	69
QY	236	TATGTTCC-----GGAGGGGATAAGCCAGCACCTGCCAACCC-----	271
DB	70	TrpProProSerSerArgGlyGlyProProGluAlaGluThrAlaGluArgAla	89
QY	272	---TCGAAGAGAAATTCGGCTCGCTCCCTCAACCGCAAGAGGGTTCGGTTTAGCAGAG	328
DB	90	GlyTrpLysThrAsnPheArgCysAlaLeuArgSerThrArgPheValMetLeuArg	109
QY	329	GACCGGAGCAAGGACCCTCAGCACCCACATAAAATCTACGAGTTGTG-----	376
DB	110	AspAsnSerGlyAspProAlaAspProHisLysValTyAlaLeuSerArgGluLeuCys	129
QY	377	---AACTCAGGAGTTGG-----GACTTTCCAGCCAGCACACCTCTCCGGAC	421
DB	130	TrpArgGluGlyProGlyThrAspGlnThrGluAlaGluAlaProAlaAlaValPro	149
QY	422	ACCAATGTTGGA-----	433
DB	150	ProGlnGlyGlyProProGlyProPheLeuAlaHisThrHisAlaGlyLeuGlnAlaPro	169
QY	433	-----	433
DB	170	GlyProLeuProAlaProAlaGlyAspLysGlyAspLeuLeuGlnAlaValGlnGln	189
QY	433	-----	433
DB	190	SerCysLeuAlaAspHisLeuLeuThrAlaSerTrpGlyAlaAspProValProThrLys	209
QY	433	-----	433
DB	210	AlaProGlyGluGlyGlnGluGlyLeuProLeuThrGlyAlaCysAlaGlyProGly	229
QY	434	-----GGCAGTACTTCTGAT	448
DB	230	LeuProAlaGlyGluLeuTyArgTrpAlaValGluThrThrProSerProThrSerAsp	249
QY	449	ACCCAGGAAGACATTCCTGGATGAGTTACTGGGTAAACATGGTGTGGCCCACTCCAGAT	508

D	b	250	ThrGlnGluAspIleLeuAspGluLeuLeuGlyAsnMetValLeuAlaProLeuProAsp	269
Q	y	509	CCGGGACCCCAAGCCTGGCTGTAGCCCTGTAGAGCCTGCCCTCAGCCCTGGGGAGCCCC	568
D	b	270	ProGlyProProSerLeuAlaValAlaProGluProCysProGlnProLeuAArgSerPro	289
Q	y	569	AGCTTGAGACAAATCCCACCTCCCTCCCAACCTGGGGCCCTCTGAGAACCCACTGAGAGCG	628
D	b	290	SerLeuAspAsnProThrProPheProAsnLeuGlyProSerGluAsnProLeuLysArg	309
Q	y	629	CTGTTGGTGGCGGGGAAGAGTGGGAGTTCGAGGTGCACAGCCCTTCACCGGGGCCCCCAA	688
D	b	310	LeuLeuValProGlyGluGluTyrGluPheGluValThrAlaPheTyrArgGlyArgGln	329
Q	y	689	GTCCTTCAGAGACCATCTCTCCCGGGAGGGCCTGGCTGTGTGGGTCCGAAGTGGGA	748
D	b	330	ValPheGlnGlnThrIleSerCysProGluGlyLeuArgLeuValGlySerGluValGly	349
Q	y	749	GACAGGACGCTCGCTGGATGCCAGTCACACTGCCAGACCTCGCATGTCCCTGACAGAC	808
D	b	350	AspArgThrLeuProGlyTyrProValThrLeuProAspProGlyMetSerLeuThrAsp	369
Q	y	809	AGGGAGTGTATGACTACGTAGGCATGTCTGAGTGCCTGGTGGGGACTGGCTCTC	868
D	b	370	ArgGlyValMetSerTyrValArgHisValLeuSerCysLeuGlyGlyLeuAlaLeu	389
Q	y	869	TGGCGGCGGGGAGTGCCTCTGGCGCCAGCGGTGGGGCAGCTGCCACACATACTGGCA	928
D	b	390	TrpArgAlaGlyGlnTyrLeuTrpAlaGlnArgLeuGlyHisCysHisThrTyrTrpAla	409
Q	y	929	GTGAGCGAGGAGCTGCTCCCAACAGCGGGCATGGGCTCATGGCGAGGTCCCAAGGAC	988
D	b	410	ValSerGluGluLeuLeuProAsnSerGlyHisGlyProAspGlyGluValProLysAsp	429
Q	y	989	AAGGAAGAGGCGTGTGTGACCTGGGGCCCTTCATTGTAGATCTGATTACCTTCACGAA	1048
D	b	430	LysGluGlyGlyValPheAspLeuGlyProPheIleValAspLeuIleThrPheThrGlu	449
Q	y	1049	GGAAACGACGCTCACACAGCTATGCCCCTCTGGTTCTGTGTGGGGAGTCATGGCCCCAG	1108
D	b	450	GlySerGlyArgSerProArgTyrAlaLeuTrpPheCysValGlyGluSerTrpProGln	469
Q	y	1109	GACACGCGGTGACCAAGAGGCTGTGTGTGTCAAGTTGTGCCCCAGCTGCCTCAGGGCC	1168
D	b	470	AspGlnProTrpThrLysArgLeuValMetValLysValValProThrCysLeuArgAla	489
Q	y	1169	TTGGTAGAAATGGCCCGGTAGGGGTGCCTCTCCCTGGAGAATACTGTGGACTGCAC	1228
D	b	490	LeuValGluMetAlaArgValGlyGlyAlaSerSerLeuGluAsnThrValAspLeuHis	509
Q	y	1229	ATTTCACACGCCACCACTCTCCCTCACTCCGACCAAGTACAGAGCCTACTCTGAGGAC	1288
D	b	510	IleAspAsnAspHisProLeuAspLeuAspAspAspGlnTyrLysAlaTyrLeuGlnAsp	529
Q	y	1289	TTGGTGGAGCGCATGGATTTCCAGGGCCCTGGGAGAGC	1327
D	b	530	LeuValGluGlyGlyAspPheGlnGlyProGlyGluSer	542
RESULT 11				
AAG00953				
ID		AAG00953 standard; protein; 102 AA.		
XX	AC	AAG00953;		
XX	DT	06-OCT-2000 (first entry)		
XX	DE	Human secreted protein, SEQ ID NO: 5034.		
XX	KW	Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;		
XX	KW	gene therapy; chromosome mapping.		
XX	OS	Homo sapiens.		

RESULT 11

RESOLUT II
AAG00953

ID AAGO

XX

AC AAG0

XX

DT
06-0
vvyDE
Huma

DE XX

KW Human

KW gene

XX

Homocysteine

XX

PN EP1033401-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 21-FEB-2000; 2000EP-00200610.
 XX
 PR 26-FEB-1999; 99US-0122487P.
 XX
 XX (GEST) GENSET.
 XX
 XX Dumas Milne Edwards J, Duclert A, Giordano J;
 PI
 XX WPI; 2000-500381/45.
 DR N-PSDB; AAC00959.
 XX
 XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures.
 XX
 PS Claim 13; SEQ ID NO 5034; 71pp + Sequence Listing; English.
 XX
 CC The present sequence is a polypeptide encoded by one of a large number of
 CC 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs were
 CC prepared from total human RNAs or polyA+ RNAs derived from 30 different
 CC tissues. EST sequences usually correspond mainly to the 3' untranslated
 CC region (UTR) of the mRNA because they are often obtained from oligo-dT
 CC primed cDNA libraries. Such ESTs are not well suited for isolating cDNA
 CC sequences derived from the 5' ends of mRNAs and even in those cases where
 CC longer cDNA sequences have been obtained, the full 5' UTR is rarely
 CC included. 5' ESTs are derived from mRNAs with intact 5' ends and can
 CC therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs
 CC are also used in diagnostic, forensic, gene therapy and chromosome
 CC mapping procedures. They are used to obtain upstream regulatory sequences
 CC and to design expression and secretion vectors
 XX
 SQ Sequence 102 AA;
 Alignment Scores:
 Pred. No.: 1,468-32 Length: 102
 Score: 546.00 Matches: 100
 Percent Similarity: 98.04% Conservatives: 0
 Best Local Similarity: 98.04% Mismatches: 2
 Query Match: 20.23% Indels: 0
 DB: 3 Gaps: 0
 US-09-975-253-1 (1-1426) x AAG00953 (1-102)
 QY 47 ATGGGAACCCCAAGCCACGNTCTGCTGGTGTGTCAGTGCAGCTGGCGGCAA 106
 Db 1 MetGlyThrProLysProArgIleLeuProTrpLeuValSerGlnLeuAspLeuGlyGln 20
 QY 107 CTGGAGGCGTGGCTGGTGAACAGAGCCGACGCGCTCCGATCCCTTGGAGACAC 166
 Db 21 LeuGluGlyValAlaTrpValAsnLysSerArgThrArgPheArgIleProTrpLysHis 40
 QY 167 GGCCTACGGCAGGATGCACAGCAGGAGATTTCGAATCTCCAGCGCTGGCCGAGGCC 226
 Db 41 GlyLeuArgGlnAspAlaGlnGlnGluAspPheGlyIlePheGlnAlaTrpAlaGluAla 60
 QY 227 ACTGCTGCATATGTTCCCGGGAGGGATAGAGCCAGACCTGCCAATCTGGAAGAGGATTC 286
 Db 61 ThrGlyAlaTyrrValProGlyArgAspLysProAspLeuProThrTrpLysArgAsnPhe 80
 QY 287 GCCTTCGCCCTCAACCGCAAGAGGTTGCGTTTAGCAGAGGACCGAGCAAGACCTT 346
 Db 81 ArgSerAlaLeuAsnArgLysGluGlyLeuArgLeuAlaGluAspArgSerLys***Pro 100
 QY 347 CACGAC 352
 Db 101 HisAsp 102
 RESULT 12
 ADD01173

ID ADD01173 standard; protein; 488 AA.
 XX
 AC ADD01173;
 XX
 DT 01-JAN-2004 (first entry)
 XX
 DE Human nucleic acid-associated protein NAAP-11 SEQ ID NO:11.
 XX
 KW human; nucleic acid-associated protein; NAAP; cytostatic;
 KW antiarteriosclerotic; anti-HIV; antiallergic; cerebroprotective;
 KW antiparkinsonian; anticonvulsant; nootropic; neuroprotective;
 KW antiinflammatory; ophthalmological; thymimetic; antiarthritic;
 KW hepatotropic; antibacterial; virucide; protozoicide; antiparasitic;
 KW fungicide; gene therapy; cell proliferative disease; cancer;
 KW atherosclerosis; hepatitis; neurological disorder; Parkinson's disease;
 KW Alzheimer's disease; stroke; epilepsy; developmental disorder;
 KW renal tubular acidosis; anaemia; glaucoma; hypothyroidism;
 KW autoimmune disorder; inflammatory disorder; AIDS; allergy;
 KW atopic dermatitis; arthritis; infection.
 XX
 OS Homo sapiens.
 XX
 PN WO2003054219-A2.
 XX
 PD 03-JUL-2003.
 XX
 PF 18-DEC-2002; 2002WO-US041115.
 XX
 PR 19-DEC-2001; 2001US-0343004P.
 PR 11-JAN-2002; 2002US-0347633P.
 PR 25-JAN-2002; 2002US-0351749P.
 PR 22-FEB-2002; 2002US-0359498P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Azimzai Y, Baughn MR, Becha SD, Borowsky ML, Burford N;
 PI Elliott VS, Emerling BW, Forsythe IJ, Gorvad AE, Griffin JA;
 PI Kabie AE, Khare R, Lal PG, Lee EA, Lee SY, Li JX, Marquis JP;
 PI Ramkumar J, Richardson TW, Sprague WW, Swarnakar A, Tang YT;
 PI Chawla NK, Warren BA, Yue H;
 DR WPI: 2003-559157/52.
 DR N-PSDB; ADD01230.
 XX
 PT New human nucleic acid-associated proteins (NAAP), useful for diagnosing,
 PT treating and preventing diseases or conditions associated with the
 PT aberrant NAAP expression, e.g. cancer, hepatitis, AIDS, atherosclerosis,
 PT infections.
 XX
 PS Claim 1; SEQ ID NO 11; 405pp; English.
 XX
 CC The present invention describes human nucleic acid-associated proteins
 CC designated NAAP-1 to NAAP-57. The human NAAPs have cytostatic,
 CC antiarteriosclerotic, anti-HIV, antiallergic, cerebroprotective,
 CC antiparkinsonian, anticonvulsant, nootropic, neuroprotective,
 CC antiinflammatory, ophthalmological, thymimetic, antiarthritic,
 CC hepatotropic, antibacterial, virucide, protozoicide, antiparasitic and
 CC fungicide activities, and can be used in gene therapy. The NAAP protein
 CC and polynucleotide sequences can be used in diagnosing, treating and
 CC preventing diseases or conditions associated with the decreased
 CC expression or overexpression of NAAP, such as cell proliferative diseases
 CC (e.g. cancer, atherosclerosis, hepatitis), neurological disorders
 CC (Parkinson's disease, Alzheimer's disease, stroke, epilepsy),
 CC developmental disorders (renal tubular acidosis, anaemia, glaucoma,
 CC hypothyroidism), autoimmune/inflammatory disorders (AIDS, allergies,
 CC atopic dermatitis, arthritis) and infections (e.g. bacterial, viral,
 CC parasitic, protozoal, fungal). The present sequence represents human NAAP
 CC -11, from the present invention.
 XX
 SQ Sequence 488 AA;
 Alignment Scores:
 Pred. No.: 3,03e-29 Length: 488

Score:	504.50	Matches:	163
Percent Similarity:	44.47%	Conservative:	62
Best Local Similarity:	32.21%	Mismatches:	186
Query Match:	18.69%	Indels:	95
DB:	7	Gaps:	22

US-09-975-253-1 (1-1426) x ADD01173 (1-488)

QY	2	GTTCAGCTGCCGACGACCCCGACCTTCCATCGTAGGCGGACCATGGAAACCCCAAG	61
Db	5	ileProValAlaProThrProProArgVal	15
QY	62	CCACGGNTCCCTGCTGGCTGTGCTGCGAGCTGGGCACTGGAGGGCGCTGGCC	121
Db	16	---ArgLeuLysProTriLeuValAlaGlnValAsnSerCysGlnTyrProGlyLeuGln	34
QY	122	TGGTGTAACAAGCGCGGCTCCGATCCCTTGAACACGCGCTACGCGAGGAT	181
Db	35	TrpValAsnGlyLysLeuLysLeuPheCysIleProTrpArgHisAlaThrArgHisGly	54
QY	182	GCACAGCAGGAG---GATTTCCGAATCTTCCAGGCTGGCGGAGCCACTGGTCATAT	238
Db	55	ProSerGlnAspGlyAspAsnThrIlePheLysAlaTrpAlaLysGluThrGlyLysTyr	74
QY	239	GTTCGGGAGGATTAAGCAGCTGCCAACCTGGAAGAGAAATTCGCTCTGCCCTC	298
Db	75	ThrGluGlyValAspGluAlaAspProAlaLysTrpLysAlaAsnLeuArgCysAlaLeu	94
QY	299	RACCGAAGAGGGTTGGTTTACGAGAGCGGAGCAAGGAC---CTCAGCAGCCCA	355
Db	95	AsnLysSerArgAspPheArgLeuIleTyrAspGlyProArgAspMetProGlnPro	114
QY	356	CATAAATCTACAGTTTGTGAATCAGGAGTTGGGACTTTCCAGCAGACACTCT	415
Db	115	TyrLysIleTyrGluValCysSerAsnGlyPro-----AlaProThrAspSerGln	131
QY	416	CCG-----GACACCAATGTGAGGAGTACTTCTGATACCCAGGAGACATCTCGAT	469
Db	132	ProProGluAspTyrSerPheGlyAlaGlyGluGluGluGluGluGluGluGln	151
QY	470	GAGTTACTGGGTAAACATG-----GTGTTGGGCCCA	499
Db	152	ArgMetLeuProSerLeuLeuThrGluAspValLysTrpProProThrLeuGlnPro	171
QY	500	CTCCAGATCCGGACCCCAAGCTGGCTGTAGCCCTGAGCCCTGAGCCCTGAGCCCTG	559
Db	172	---ProThrLeuGlnProProValValLeuGlyProProAlaProAspProSerProLeu	190
QY	560	CGGAGCCCGAGCTGGACATCCACTCCCTCCCAACCTGGGGCCCTCTGAG-----	613
Db	191	---AlaProProProGlyAsnProAlaGlyPheArgGluLeu---LeuSerGluValLeu	208
QY	614	-----AACCCACTGAAGCGCTGTGGTCCG---GGGGAAGAG-----	649
Db	209	GluProGlyProLeuProAlaSerLeuProProAlaGlyGluGlnLeuLeuProAspLeu	228
QY	650	-----TGGAGTTCCAGGTGACAGCTTCTACCGGGC	682
Db	229	LeuIleSerProHisMetLeuProLeuThrAspLeuGluIleLysPheGlnTyrArgGly	248
QY	683	CGCAAGTCTCCAGCAGACATCTCTGCCGAGGGCTGCGGCTGGTGGGTCGGA	742
Db	249	ArgProProArgAlaLeuThrIleSerAsnProHisGlyCysArgLeuPheTyrSerGln	268
QY	743	GTG-----GGAGACAGCAGCTGCTCGCTGGATGGCCA	772
Db	269	LeuGluAlaThrGlnGluGlnValGluLeuPheGlyProIleSerLeuGlu-----Gln	286
QY	773	GTCACTCCAGACCTGGGCTGTCTCTGACAGAGGGAGTGTATGAGTACTGAGG	832
Db	287	ValArgPheProSerProGluAspIleProSerAspLys---GlnArgPheTyrThrAsn	305
QY	833	CATGTGCTGAGCTGCTGGTGGGGAGTGGCTCTCTCGGCGGCGGCGAGTGGCTCTGG	892

Db	306	GlnLeuLeuAspValLeuAspArgGlyLeuLeuGlnLeuGlnGlnGlnGlnAspLeuTyr	325
QY	893	GCCAGGCGGCTGGGCTGCTCCACACATCTGG-----GCAGTGGAGGAGG	940
Db	326	AlaIleArgLeuCysGlnCysLysValPheTrpSerGlyProCysAlaSerAlaHisAsp	345
QY	941	CTGCTCCCAACAGCGGCTGCTGATGGCGAGTCCCAAGGACCAAGCAAGAGGAGGC	1000
Db	346	SerCysProAsn-----ProIleGlnArgGluValLysThrLys	358
QY	1001	GTGTTTACCTGGGCGCTTCACTTGTAGTCTGATCTTACCGAAGGA---AGCGGA	1057
Db	359	LeuPheSerLeuGluHisPheLeuAsnGluLeuIleLeuPheGlnLysGlyGlnThrAsn	378
QY	1058	CGCTCACACGCTATGCCCTCTGGTTCTGTGGGGAGTCTATGCCCGCAGGACCGCG	1117
Db	379	ThrProProProPheGluIlePhePheCysPheGlyGluGluTrpProAspArgLysPro	398
QY	1118	TGGACCAAGAGCTGCTGATGTCAAGTTGTGCCACGTGCTCAGGCGCTTGTGTAGAA	1177
Db	399	ArgGluLysLysLeuIleThrValGlnValProValAlaAlaAlaArgLeuLeuGlu	418
QY	1178	ATGCGCGGCTAGGGGCTGCTCTCCCTGGAGAACTGTGGAGCTGCACATTTCCAAAC	1237
Db	419	MetPheSer---GlyGluLeuSerTrpSerAlaAspSerIleArgLeuGlnIleSerAsn	437
QY	1238	AGCCA-----CCCACTCTCCCTCAC-----	1257
Db	438	---ProAspLeuLysAspArgMetValGluGlnPheLysGluLeuHisIleTrpGlnS	457
QY	1258	-----CTCCGACAGTACAGGCTTACCTGCGAGACTTGTGGAGGCGCATGGAT	1306
Db	457	erGlnGlnArgLeuGlnProValAlaGlnAlaProGlyAlaGlyLeuGlyValGly-	476
QY	1307	TTCCAGGCGCTGG 1320	
Db	477	--GlnGlyProTrp 480	

RESULT 13
ADJ75465
ID ADJ75465 standard; protein; 516 AA.
XX AC ADJ75465;
XX DT 20-MAY-2004 (first entry)
XX DE
XX Marker gene related amino acid sequence SEQ ID NO:717.
XX bronchial asthma; chronic obstructive pulmonary disease;
XX respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;
XX gene therapy; marker.
XX Homo sapiens.
XX EPI394274-A2.
XX 03-MAR-2004.
XX 04-AUG-2003; 2003EP-00254857.
XX 06-AUG-2002; 2002JP-00229312.
XX 20-MAR-2003; 2003JP-00077212.
XX (GENO-) GENOX RES INC.
XX Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuhara K;
XX WPI; 2004-193155/19.
XX Testing for bronchial asthma or chronic obstructive pulmonary disease by
XX comparing the expression level of a marker gene in a biological sample
XX from a subject with the expression level of the gene in a sample from a

healthy subject.

Example 11; SEQ ID NO 717; 241pp; English.

The present invention describes a method of testing for bronchial asthma or chronic obstructive pulmonary disease. The method comprises determining the expression level of a marker gene in a biological sample from a subject, comparing the expression level determined with the expression level of the marker gene in a biological sample from a healthy subject, and judging whether the subject has bronchial asthma or chronic obstructive pulmonary disease. The marker gene comprises: (a) a group of genes (S1) whose expression levels increase when respiratory epithelial cells are stimulated with interleukin-13; or (b) a group of genes (S2) whose expression levels decrease when respiratory epithelial cells are stimulated with interleukin-13. Also described: (1) a reagent (I) for testing for bronchial asthma or chronic obstructive pulmonary disease; (2) a kit for screening for a candidate compound for a therapeutic agent to treat bronchial asthma or chronic obstructive pulmonary disease; (3) an animal model for bronchial asthma or chronic obstructive pulmonary disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a method for producing an animal model for bronchial asthma or chronic obstructive pulmonary disease; (6) a therapeutic agent for bronchial asthma or chronic obstructive pulmonary disease, comprising the compound, a marker gene or an antisense nucleic acid corresponding to a portion of the marker gene, a ribozyme, a polynucleotide that suppresses the expression of the gene through an RNAi effect or an antibody recognising a protein encoded by a marker gene; and (7) a DNA chip for testing for bronchial asthma or a chronic obstructive pulmonary disease, on which a probe has been immobilised to assay a marker gene. (I) has respiratory and antiallergic activities, and can be used in gene therapy. The method is useful for testing for or screening for a therapeutic agent for bronchial asthma or chronic obstructive pulmonary disease. The present sequence is used in the exemplification of the present invention.

Sequence 516 AA;

Alignment Scores:
 Pred. No.: 1-27e-25 Length: 516
 Score: 456.50 Matches: 145
 Percent Similarity: 40.15% Conservative: 65
 Best Local Similarity: 27.72% Mismatches: 202
 Query Match: 16.91% Indels: 111
 DB: 8 Gaps: 20

US-09-975-253-1 (1-1426) x ADU75465 (1-516)

PT healthy subject.
 XX Example 11; SEQ ID NO 717; 241pp; English.
 XX The present invention describes a method of testing for bronchial asthma or chronic obstructive pulmonary disease. The method comprises determining the expression level of a marker gene in a biological sample from a subject, comparing the expression level determined with the expression level of the marker gene in a biological sample from a healthy subject, and judging whether the subject has bronchial asthma or chronic obstructive pulmonary disease. The marker gene comprises: (a) a group of genes (S1) whose expression levels increase when respiratory epithelial cells are stimulated with interleukin-13; or (b) a group of genes (S2) whose expression levels decrease when respiratory epithelial cells are stimulated with interleukin-13. Also described: (1) a reagent (I) for testing for bronchial asthma or chronic obstructive pulmonary disease; (2) a kit for screening for a candidate compound for a therapeutic agent to treat bronchial asthma or chronic obstructive pulmonary disease; (3) an animal model for bronchial asthma or chronic obstructive pulmonary disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a method for producing an animal model for bronchial asthma or chronic obstructive pulmonary disease; (6) a therapeutic agent for bronchial asthma or chronic obstructive pulmonary disease, comprising the compound, a marker gene or an antisense nucleic acid corresponding to a portion of the marker gene, a ribozyme, a polynucleotide that suppresses the expression of the gene through an RNAi effect or an antibody recognising a protein encoded by a marker gene; and (7) a DNA chip for testing for bronchial asthma or a chronic obstructive pulmonary disease, on which a probe has been immobilised to assay a marker gene. (I) has respiratory and antiallergic activities, and can be used in gene therapy. The method is useful for testing for or screening for a therapeutic agent for bronchial asthma or chronic obstructive pulmonary disease. The present sequence is used in the exemplification of the present invention.
 XX Sequence 516 AA;
 SQ Alignment Scores:
 Pred. No.: 1-27e-25 Length: 516
 Score: 456.50 Matches: 145
 Percent Similarity: 40.15% Conservative: 65
 Best Local Similarity: 27.72% Mismatches: 202
 Query Match: 16.91% Indels: 111
 DB: 8 Gaps: 20
 US-09-975-253-1 (1-1426) x ADU75465 (1-516)
 QY 5 CCAGCTGCC-----CGCAGCGCCCGACCTTCATCGTAGCGCGACCATGGGAAACCCCA 58
 DB 7 ProAlaAlaGlyProAspSerProArgPro-----GlyThrArgAlaAlaPro 23
 QY 59 AGCCACGGNTCTGCTCCCTGGTGTGTGCGAGCTGGAGCTGGGCAACTGGAGGGCGTG 118
 DB 24 ArgValLeuPheGlyGluTrpLeuLeuGlyGluLeuSerSerGlyCysTyrGluGlyLeu 43
 QY 119 GCCTGGGTGAACAAGACGCGCGCTTCGCGATCCCTTGGAGCAGCGCCCTAGCGCAG 178
 DB 44 GlnTrpLeuAspGluAlaArgThrCysPheArgValProTrpPheHisPheAlaArgLys 63
 QY 179 GATGACAGCAGGAGATTTCGGAATCTTCCAGGCTGGCGCGAGCCACTGGTGCATAT 238
 DB 64 AspLeuSerGluAlaAspAlaArgilePheLysAlaTrpAlaValAlaArgGlyArgTrp 83
 QY 239 GTTCCC-----GGGAGGATAGCCAGACCTGCCAAC----- 271
 DB 84 ProProSerSerArgGlyGlyGlyProProGluAlaGluThrAlaGluAlaGly 103
 QY 272 TGAAGAGAAATTCGCTCTCCCTCAACCCAAAGAGAGGTTCGTTTACACAGGAC 331
 DB 104 TrpLysThrAsnPhaArgCysAlaLeuArgSerThrArgPheValMetLeuArgAsp 123
 QY 332 CGGACAAAGGACCTCAGACCCACATATAATCTACGAGTTTGTG----- 376
 DB 124 AsnSerGlyAspProAlaAspProHisLysValTyrAlaLeuSerArgGluLeuCysTrp 143

QY 377 AACTCAGGAGTTGG-----GACTTTTCCAGCCAGACACCTCTCCGACACC 424
 DB 144 ArgGluGlyProGlyThrAspGlnThrGluAlaGluAlaProAlaAlaValProPro 163
 QY 425 AATGCTGGA----- 433
 DB 164 GlnGlyGlyProGlyProGlyProPheLeuAlaHisThrHisAlaGlyLeuGlnAlaProGly 183
 QY 434 -----GSCAGTACTCTGAT-----ACCCAGGAA 457
 DB 184 ProLeuProAlaProAlaGlyAspGluGlyAspLeuLeuGlnAlaValGlnGlnSer 203
 QY 458 GACATCTCGATGAGTTACTG---GGTAACATGGTGTGGCCCTCCCTCCCA 505
 DB 204 CysLeuAlaAspHisLeuLeuThrAlaSerTyrGlyAlaAspProValProThrLysAla 223
 QY 506 -----GATCCGGGACCC 517
 DB 224 ProGlyGluGlnGluGlyLeuProLeuThrGlyAlaCysAlaGlyGlyProGlyLeu 243
 QY 518 CCA-----AGCCTGGCTGTAGCCCTGAGCCCTGAGCCCTCCCTAGCC 556
 DB 244 ProAlaGlyGluLeuTyrGlyTrpAlaValGluThrProSerProGlyProGlnPro 263
 QY 557 CTGGGAGCCCGAGCTTGGCAATCCCACTCCC-----TTCCCAAC 598
 DB 264 AlaAlaLeuThrThrGlyGluAlaAlaAlaProGluSerProHisGlnAlaGluProTyr 283
 QY 599 CTGGGCGCTCTGAGAACCCACTGAGCGGTGTGTGTCGCGGGGAAAGTGGAGTTC 658
 DB 284 LeuSerProSerProSerAlaCysThrAlaValGlnGluPro---SerProGlyAlaLeu 302
 QY 659 GAGGTGACAGCCTTCTACCGGGCGCCCAAGTCTTCCAGACGACCATCTCTGCTCCCGGAG 718
 DB 303 AspValThrIleMetTyrLysGlyArgThrValLeuGlnLysValValGlyHisProSer 322
 QY 719 GGCCTCGCGCTGGTGGG---TCCGAAGTGGAGACAGGACGCTGCTGGATGGCCAGTC 775
 DB 323 CysThrPheLeuTyrGlyProProAspProAlaValAlaThrAspProGlnGlnVal 342
 QY 776 ACCTGTCAGACCCCTGGCATGTCCCTGACAGACAGGGAGTGTATGAGTACGTAGCAT 835
 DB 343 AlaPheProSerProAla---GluLeuProAspGlnLysGlnLeuArgTyrThrGluGlu 361
 QY 836 GTGCTGAGCTGCTGGTGGGAGCTGCTCTGCGCGCGCGGAGGAGCTGCTCCCAACAGC 895
 DB 362 LeuLeuArgHisValAlaProGlyLeuHisGluLeuArgGlyProGlnLeuTrpAla 381
 QY 896 CAGCGCTGGGCACTGCCACACATCTGGGCGAGTGAGCGAGGAGCTGCTCCCAACAGC 955
 DB 382 ArgArgMetGlyLysCysLysValTyrTrpGluValGly----- 394
 QY 956 GGCATGGCGCTGATGGGAG-----GTCCCCAAGGAC 988
 DB 395 -----GlyProProGlySerAlaSerProSerThrProAlaCysLeuLeuProArgAsn 412
 QY 989 AAGGAAGGAGCGCTTTGACCTGGGCGCTTCTGATGATCTGATACCTTACCTTACCGAA 1048
 DB 413 CysAspThrProIlePheAspPheArgValPhePheArgGluLeuValGluPheArgAla 432
 QY 1049 GGAAGCGGACGC---TCACACGCTATCCCTCTGTGTGGGGAGTGCATGGGCC 1105
 DB 433 ArgGlnArgArgGlySerProArgTyrThrIleTyrLeuGlyPheGlyGlnAspLeuSer 452
 QY 1106 CAGGACGAGCGTGGACCAAGAGCTCTGATGTGTCAGGTGTGCGCCAGTGCCTCAGG 1165
 DB 453 AlaGlyArgProLysGluLysSerLeuValLeuValLysLeuGluProTrpLeuCysArg 472
 QY 1166 GCCTTGGTAGAAATGGCCCGGAGGGGTGCTCTCCCTCCCTGGAG---AATACTGTGAC 1222
 DB 473 ValHisLeuGluGlyThrGlnArgGluGlyValSerSerLeuAspSerSerSerLeuSer 492

AC	ADJ75413;
XX	
DT	20-MAY-2004 (first entry)
XX	
DE	Marker gene related amino acid sequence SEQ ID NO:665.
XX	
KW	bronchial asthma; chronic obstructive pulmonary disease;
XX	
KW	respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;
XX	
KW	gene therapy; marker.
XX	
OS	Homo sapiens.
XX	
PN	EP1394274-A2.
XX	
PD	03-MAR-2004.
XX	
PF	04-AUG-2003; 2003EP-00254857.
XX	
PR	06-AUG-2002; 2002JP-00229312.
XX	
PR	20-MAR-2003; 2003JP-00077212.
XX	
PA	(GENO-) GENOX RES INC.
XX	
PI	Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuwara K;
XX	
PI	WPI; 2004-193155/19.
XX	
DR	
XX	Testing for bronchial asthma or chronic obstructive pulmonary disease by
PT	comparing the expression level of a marker gene in a biological sample
XX	
PT	from a subject with the expression level of the gene in a sample from a
XX	
PT	healthy subject.
XX	
XX	Example 11; SEQ ID NO 665; 241pp; English.
BS	
XX	The present invention describes a method of testing for bronchial asthma
CC	or chronic obstructive pulmonary disease. The method comprises
CC	determining the expression level of a marker gene in a biological sample
CC	from a subject, comparing the expression level determined with the
CC	expression level of the marker gene in a biological sample from a healthy
CC	subject, and judging whether the subject has bronchial asthma or chronic
CC	obstructive pulmonary disease. The marker gene comprises: (a) a group of
CC	genes (S1) whose expression levels increase when respiratory epithelial
CC	cells are stimulated with interleukin-13; or (b) a group of genes (S2)
CC	whose expression levels decrease when respiratory epithelial cells are
CC	stimulated with interleukin-13. Also described: (1) a reagent (I) for
CC	testing for bronchial asthma or chronic obstructive pulmonary disease;
CC	(2) a kit for screening for a candidate compound for a therapeutic agent
CC	to treat bronchial asthma or chronic obstructive pulmonary disease; (3)
CC	an animal model for bronchial asthma or chronic obstructive pulmonary
CC	disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a
CC	method for producing an animal model for bronchial asthma or chronic
CC	obstructive pulmonary disease; (6) a therapeutic agent for bronchial
CC	asthma or chronic obstructive pulmonary disease, comprising the compound,
CC	a marker gene or an antisense nucleic acid corresponding to a portion of
CC	the marker gene, a ribozyme, a polynucleotide that suppresses the
CC	expression of the gene through an RNAi effect or an antibody recognising
CC	a protein encoded by a marker gene; and (7) a DNA chip for testing for
CC	bronchial asthma or a chronic obstructive pulmonary disease, on which a
CC	probe has been immobilised to assay a marker gene. (I) has respiratory
CC	and antiasthmatic activities, and can be used in gene therapy. The method
CC	for screening for a therapeutic agent for

Fri Oct 29 11:11:43 2004

493 LeuCysLeuSerSerAlaAsnSerLeu-----TyrAspAspIleGluCysPheLeu 509

1283 CAGGACTTG 1291
510 MetGluLeu 512

Search completed: October 28, 2004, 15:22:21
Job time : 242 secs

377 AACTCAGAGTTGGG-----GACTTTTCCCGACGACACACTCTCCGACACC 424
144 ArgGluGlyProGlyThrAspGlnThrGluAlaGluAlaProAlaValProPro 163
425 AATGTTGA----- 433
164 GlnGlyGlyProProGlyProPheLeuAlaHisThrHisAlaGlyLeuGlnAlaProGly 183
434 -----GGCAGTACTTCTGAT-----ACCCAGGAA 457
184 ProLeuProAlaProAlaGlyAspGluGlyAspLeuLeuGlnAlaValGlnGlnSer 203
458 GACATTCTGATGAGTTACTG---GGTAACATGGTGTGGCCCACTCCCA----- 505
204 CysLeuAlaAspHisLeuLeuThrAlaSerTrpGlyAlaAspProValProThrIysAla 223
506 -----GATCCGGGACCC 517
224 ProGlyGluGlyGlnGluGlyLeuProLeuThrGlyAlaCysAlaGlyProGlyLeu 243
518 CCA-----AGCTGGGTGTAGCCCTGAGCCCTGAGCCCTGAGCCCTCAGCCC 556
244 ProAlaGlyGluLeuTyrglyTrpAlaValGluThrProSerProGlyProGlnPro 263
557 CTGCGAGCCCGAGCTTGGACAAATCCCACTCCC-----TTCCCAAAC 598
264 AlaAlaLeuThrThrGlyGluAlaAlaAlaProGluSerProHisGlnAlaGluProTyrr 283
599 CTGGGCGCTCTGAGAACCCACTGAAGCGCTGTGGTCCGGGGAAGAGTGGGAGTTC 658
284 LeuSerProSerProAlaCysThrAlaValGlnGluPro---SerProGlyAlaLeu 302
659 GAGGTGACAGCTTCTACCGGCGCGCAAGTCTTCCAGCAGACCATCTCTCCCGCGGAG 718
303 AspValThrIleMetTyrllysGlyArgThrValLeuGlnLysValValGlyHisProSer 322
719 GGCTCTGGCTGTGGGG---TCCGAAGTGGGAGACAGGACGCTGCTGATGGCCAGTC 775
323 CysThrPheLeuTyrglyProProAlaValArgAlaThrAspProGlnGlnVal 342
776 ACCTGCGACAGCTGTCATGTCCTCAGACAGAGGGAGTGATGAGCTACGTGAGGCAT 835
343 AlaPheProSerProAla---GluLeuProAspGlnLysGlnLeuArgTyrrThrGlu 361
836 GTGCTGAGCTGCTGGGTGGGGACTGCTCTCGGCGGCGCGGCGAGTGGCTCTGGCC 895
362 LeuLeuArgHisValAlaProGlyLeuHisLeuGluLeuArgGlyProGlnLeuTrpAla 381
896 CAGCGCTGGGCGCACTGCCACACATCTGGCGAGTGAGCGGAGAGTGTCTCCCAACAGC 955
382 ArgArgMetGlyLysCysLysValTyrrTrpGluValGly----- 394
956 GGGCATGGCTGTATGGCGAG-----GTCCCAAGGAC 988
395 -----GlyProProGlySerAlaSerProSerThrProAlaCysLeuLeuProArgAsn 412
989 AAGGAAGGAGCGGTGTGTGACTGGGCGCTTTCATTTGATCATTCATTCACGGAA 1048
413 CysAspThrProIlePheAspPheArgValPhePheArgGluLeuValGluPheArgAla 432
1049 GGAAGCGGACGC---TCACACGCTATGCTCTGCTGTGTGGGGAGTCATGGCCC 1105
433 ArgGlnArgArgGlySerProArgTyrrThrIleTyrrLeuGlyPheGlyGlnAspLeuSer 452
1106 CAGGACCGCGTGGACCAAGAGGCTCGTGTGTGTCAGAGTTGTGCCACGCTGCCTCAGG 1165
453 AlaGlyArgProLysGluLysSerLeuValLeuValLysLeuGluProTripleuCysArg 472
1166 GCCTTGGTAGAAATGGCGGTAGGGGTGCTCTCTCCCTGGAG---AATACTGTGGAC 1222
473 ValHisLeuGluGlyThrGlnArgGluGlyValSerSerLeuAspSerSerLeuSer 492
1223 CTGCACATTTCCAAACGCCACCCACTCTCCCTCACCTCCGACCAAGTACAGGCCTACCTG 1282

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	499.5	18.5	504	2	G02474		interferon regulat
2	457.5	17.0	491	2	S56753		interferon regulat
3	438.5	16.2	459	2	JC5620		interferon regulat
3	438.5	14.4	450	2	S59837		lymphoid-specific
5	351	13.0	425	2	A45064		interferon consens
6	343.5	12.7	424	2	A35861		interferon consens
7	301.5	11.2	393	2	A45017		transcription fact
8	276.5	10.2	203	2	S57836		lymphoid-specific
9	240	8.9	325	2	I52998		interferon regulat
10	235.5	8.7	399	2	JC4592		transcription fact
11	232	8.6	325	2	B31593		interferon regulat
C 12	226.5	8.5	1585	2	T31611		hypothetical prote
C 13	219	8.1	349	2	A53340		interferon regulat
C 14	215	8.0	574	2	S13301		collagen alpha 1(X

Db 35 TrpValAsnGlyGluLysLeuPheCysIleProTrpArgHisAlaThrArgHisGly 54
QY 182 GCACAGCAGGAG- -GATTTCGGAATCTTCAGGCGCTGGCGAGGCCACTGGTCATAT 238
Db 55 ProSerGlnAspGlyAspAsnThrIlePheLysAlaTrpAlaLysGluThrGlyLysTyr 74
QY 239 GTTCCCGGAGGAGTAAGACAGCAGCTGCCAACCTCGAAGAGAAATTCGGCTCTGCCCTC 298
Db 75 ThrGluGlyValAspGluAlaAspProAlaLysTrpLysAlaAsnLeuArgCysAlaLeu 94
QY 299 AACCCCAAGAGGTTGGTTTGTAGCAGAGGACCGGACGAGC- -CCTCAGCAGCCA 355
Db 95 AsnLysSerArgAspPheArgLeuIleLysAspGlyProArgAspMetProGlnPro 114
QY 356 CATAAATCTACAGATTGTGAATCAGGAGTTGGGACTTTTCCAGCCAGACACTCT 415
Db 115 TyrLysIleTyrGluValCysSerAsnGlyPro- - -AlaProThrAspSerGln 131
QY 416 CCG- - -GACACCAATGGTGAGGACAGTCTCTGTATACCCAGGACACATTCGGAT 469
Db 132 ProProGluAspTyrSerPheGlyAlaGlyGluGluGluGluGluGluGln 151
QY 470 GAGTTACTGGTAACATGGTTGGCCCTCCAGATCCGGA- - - 514
Db 152 ArgMetLeuProSerLeuSerLeuThrAspAlaValGlnSerGlyProHisMetThrPro 171
QY 514 - - - - - 514
Db 172 TyrSerLeuLeuLysGluAspValLysTyrProProThrLeuGlnProProThrLeuGln 191
QY 515 CCCCCAAGCGTGTAGCCCTGAGCCCTGAGCCCTGAGCCCTGAGCCCTGAGCCCTGAG 574
Db 192 ProProValValLeuGlyProProAlaProAspProSerProLeu- - -AlaProProPro 210
QY 575 GACAATCCCACTCCCTCCCAACCTGGGCGCTCTGAG- - -AACCCACTG 622
Db 211 GlyAsnProAlaGlyPheArgGluLeu- - -LeuSerGluValLeuGluProGlyProLeu 229
QY 623 AAGCGGTGTGTGTGCGG- - -GGGGAAGAG- - - 649
Db 230 ProAlaSerLeuProProAlaGlyGluGlnLeuLeuProAspLeuLeuLeuSerProHis 249
QY 650 - - - - -TGGGAGTTCGAGTGACAGCCTTCTACCGGCGCGCCAGCTCTCCAG 697
Db 250 MetLeuProLeuThrAspLeuGluIleLysPheGlnTyrArgGlyArgProProArgAla 269
QY 698 CAGACCATCTCTGCGCGGCGGCTGCGGTGTGGGTCCGAGTG- - - 745
Db 270 LeuThrIleSerAsnProHisGlyCysArgLeuPheTyrSerGlnLeuGluAlaThrGln 289
QY 746 - - - - -GGAGACAGACGCTGCTGGATGGCGAGTCACACTGCCAGAC 787
Db 290 GluGlnValGluLeuPheGlyProIleSerLeuGlu- - -GlnValArgPheProSer 307
QY 788 CTGGCGATGCTCCACAGACAGGAGTGTATGATGATGATGATGATGATGATGATGATGAT 847
Db 308 ProGluAspIleProSerAspLys- - -GlnArgPheTyrThrAsnGlnLeuLeuAspVal 326
QY 848 CTGGGTGGGAGTGGTCTCTGCGGCGCGGAGTGGTCTGCGGCGCGGAGTGGTCTGCGG 907
Db 327 LeuAspArgGlyLeuIleLeuGlnLeuGlnGlyGlnAspLeuTyrAlaIleArgLeuCys 346
QY 908 CACTGCCACACATACTGG- - -GCATGAGCGGAGGAGTGTCTCCCAACAGC 955
Db 347 GlnCysLysValPheTrpSerGlyProCysAlaSerAlaHisAspSerCysProAsn- - - 365
QY 956 GGGCATGGCGCTGATGGCGAGTCCCAAGGACAGGAGGAGCGGTGTGTGCTGCTGGG 1015
Db 366 - - - - -ProIleGlnArgGluValLysThrLysLeuPheSerLeuGlu 379
QY 1016 CCCTTTCATTGTATCTGATTACCTTTCACGGAAGA- - -AGCGGACGCTCACCACGCTAT 1072
Db 380 HisPheLeuAsnGluLeuIleLeuPheGlnLysGlyGlnThrAsnThrProProPhe 399

QY 1073 GCCCTCTGTTCTGTGGGAGTCAATGGCCCGCAGGACCGCTGGACCAAGAGGCTC 1132
Db 400 GluIlePhePheCysPheGlyGluTrpProAspArgLysProArgLysLysLeu 419
QY 1133 GTGATGGTCAAGGTTGTGCTCCACAGCTGCTCAGGGCTTGGTAGAAATGCGCCGGTAGGG 1192
Db 420 IleThrValGlnValProValAlaAlaArgLeuLeuLeuMetPheSer- - -Gly 438
QY 1193 GGTGCTCTCTCCTCGGAGAACTGTGACCTGACCTGCACATTCACAGACCA- - - 1242
Db 439 GluLeuSerTrpSerAlaAspSerIleArgLeuGlnIleSerAsn- - -ProAspLeuLysA 458
QY 1243 - - - - -CCCACTCTCCCTCAC- - -CTCC 1261
Db 458 spArgMetValGluGlnPheLysGluLeuHisIleTrpGlnSerGlnGlnArgLeuG 478
QY 1262 GACCACTACAGGCTTACCTGAGGACTTGTGGAGGCGATGATTTCCAGGCGCTGG 1320
Db 478 InProValAlaGlnAlaProProGlyAlaGlyLeuGlyValGly- - -GlnGlyProTrp 496
RESULT 2
S56753
interferon regulatory factor 3 - chicken
C:Species: Gallus gallus (chicken)
C:Date: 10-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C:Accession: S56753
R:Grant, C.E.; Vasa, M.Z.; Deeley, R.G.
Nucleic Acids Res. 23, 2137-2146, 1995
A:Title: cIRF-3, a new member of the interferon regulatory factor (IRF) family that is
A:Reference number: S56753; MUID:95334365; PMID:7541908
A:Accession: S56753
A>Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-491 <RA>
A:Cross-references: UNIPROT:Q90643; EMBL:U20338; NID:g790580; PIDN:AAA86995.1; PID:g790580
C:Superfamily: lymphoid-specific interferon regulatory factor
Alignment Scores:
Pred. No.: 1,18e-21 Length: 491
Score: 457.50 Matches: 146
Percent Similarity: 41.57% Conservative: 61
Best Local Similarity: 29.32% Mismatches: 174
Query Match: 16.95% Indels: 117
DB: 2 Gaps: 19
US-09-975-253-1 (1-1426) x S56753 (1-491)
QY 59 AAGCCACGGNTCCCTGCTGCTGGTGTGTCAGCTGACCTGGGGCAACTGGAGGCGTG 118
Db 12 LysLeuArgPheGlyProTrpLeuLeuAsnAlaValSerSerGlyLeuTyrArgGlyLeu 31
QY 119 GCCTGGGTGAACAGAGCCGACGCGCTTCGGATCCCTTGGAGCAGCAGCCCTACGCGCAG 178
Db 32 CysTrpIleAspProAspArgGliePheArgIleProTrpLysHisAsnAlaArgLys 51
QY 179 GATGCACAGCAGGAGGATTCGGAATCTCCAGCCTGGCGCCGAGGCACTGGTGCATAT 238
Db 52 AspValThrSerSerAspValGluIlePheLysAlaTrpAlaLysAlaSerGlyArgTyr 71
QY 239 GTTCCCGGAGGATTAAGCCAGACCTGCCAACCTCGAAGAGGAATTTCCGCTCTGCCCTC 298
Db 72 GluGlyAsnAlaGluAspPro- - -AlaLysTrpLysThrAsnPheArgCysAlaLeu 89
QY 299 AACCGCAAGAGGTTGCGTTTAGCAGAGGACCGGAGCAGGACCTCAGACCCACAT 358
Db 90 ArgSerThrHisMetPheMetLeuLeuGluAspArgSerLysCysAsnAspAspProHis 109
QY 359 AAAATCTACAGTTTGTGAATCTCAGGAGTT- - - 388
Db 110 LysValTyrAla- - -ValAlaSerGlyValProAsnAspArgGlySerGlyProVal 128
QY 389 - - - - -GGGACTTTTCCAGCCA 406

Db	129	AlaGlyAlaLeuGlnGlnProGlnLeuLeuLeuAsnHisAspLeuAlaLeuGlu	148
Qy	407	GACACCTCTCCGGAC---ACCAATGGTGGAGGCAGTACTTCTGATGACACAGAAACAT	463
Db	149	AsnThrProThrAspSerThrGluGlyValAlaAlaAlaLeuThrGlnValAspLeu	168
Qy	463	-----	463
Db	169	AspLeuLeuGlnSerValLeuGlnHisCysAsnIleSerAlaLeuGlySerGlnProThr	188
Qy	464	-----CTGGATGAGTTACTGGGTAACTGGTGTGGCCCA-----	499
Db	189	LeuTrpAlaHisThrGlyAspAlaLeuProGluAspAlaLeuLeuLeuProGlyClnAsp	208
Qy	500	-----CTCCAGATCCG-----	511
Db	209	GlyCysLeuProGlyProGlnPheGlnAspTrpArgGlnLeuGluProLeuLeuLeu	228
Qy	512	-----GGACCCCAAGCCTGCTGTA	532
Db	229	GlyAsnGlnProLeuThrGlyGlyCysGlyGlnAspGlyAlaGlyAlaLeuProVal	248
Qy	533	GCCCTCAGCCCTCCCTCAGCCCTCGGAGCCCGCTGGAC-----	577
Db	249	Ser---GluGlnCysAlaIleProAlaProSerProAlaGluGluLeuLeuPheGlnSer	267
Qy	578	---AATCCCATCTCCCTTCCAAACCTGGGGCCCTCTGAGAACCCACTGAAGCGGTGTG	634
Db	268	AlaAsnProAlaProProProAlaGlyAsp-----	278
Qy	635	GTCCCGGGGAAGTGGAGTTCGAGGTGACAGCTTCTACGGGGCGGCAAGTCTTC	694
Db	279	IleGlyGlyLeuProProLeuLeuAspIleThrIleTyrTrpArgGlyLysMetValTyr	298
Qy	695	CAGCAGACCATCTCC-----TGCCCGAGGGCCTGCGGCTGGTGGGTCCGAAGTG	745
Db	299	GlnGlnValAspSerArgCysValLeuAlaTyrGlnProLeuAspProAlaVal	318
Qy	746	GGAGACAGGACGTGCCTGGATGCCAGTCACTGCCAGACCTGGCATCTCCCTGACA	805
Db	319	AlaGlnGlnArgLeu-----ValLeuPheProSerProAla---SerLeuPro	333
Qy	806	GACAGGGAGTGATGAGTACGTGAGGACANGTGTGAGCTGCCTGGTGGGGAGTGCST	865
Db	334	AspProArgGlnArgTyrTrpGluAspLeuLeuGluVal-----AlaGlyLeuArg	351
Qy	866	CTC---TGGCGGGCGGGCAGTGCSTCTGGGCCAGCGGTGGGGCTGCCACACATAC	922
Db	352	LeuGluGlnArgAlaGlyGln---LeuLeuAlaThrArgLeuTyrLysCysLysValPhe	370
Qy	923	TGGCAGCTGAGCAGGAGCTGCTCCCAACAGCGGGCATGGGCTGATGGGAGGTCCCC	982
Db	371	TrpAlaLeuSerGlnGlnLeu-----GluGlyGlyGluProProLeuAsnLeuHis	388
Qy	983	AAGACAAGGAAGAGCGGTGTTTGAACCTGGGGCCCTTCATTGTAGATCTGATTACCTTC	1042
Db	389	ArgAspGlnGluThrThrIlePheAspPheArgValPheCysThrGluLeuArgAspPhe	408
Qy	1043	ACGGAAGA---ACGGACGCTACACAGCTATGCCCTCTGGTCTCTGTGGGGAGTCA	1099
Db	409	ArgAspSerArgArgGluArgSerProAspPheThrIlePheLeuCysPheGlyGlnCys	428
Qy	1100	TGGCCCCCAGGACACGCGTGGACCAAGAGGCTCGTGTATGGTCAAGTGTGCCCCACGTGC	1159
Db	429	PheSerSerThrLysProLysGluSerLysLeuLeuValLeuValProGlnPhe	448
Qy	1160	CTCAGGCGCTGTAGAAATGGCCCGGGTAGGGGTGCCTCTCCCTGGAGAAAT---ACT	1216
Db	449	CysGluTyrTrpTyrGluGlnValGlnArgGlyAlaSerSerLeuAsnSerGlyAsn	468
Qy	1217	GTGACCTGCACATTTCAACAGCCACCCACTCTCCCTCACTCCGACCAAGTAC	1270

```

Db 204 ProGluMetTrpIleSerSerLeuProMetThr 214
QY 653 GAGTTCGAGGTGACAGCTTCTACCGGGCGCCCAAGTCTTCCAG---CAGACCATCTCC 709
Db 215 AspLeuGluIleGlnPheTyrTyrArgGlyLysGluMetGlyGlnThrMetThrValSer 234
QY 710 TGCCCGAGGCGCTCGGCTGGTGGGTCGGAAGTGGGAGACAGACGCTGCCTGGA--- 766
Db 235 AsnProGlnGlyCysArgLeuPheTyrGlyAspLeuGly-----ProMetProAsnGln 252
QY 767 -----TGCCAGTACACTGCCAGAC-----CCNGGATG---TCCCTG 802
Db 253 GluGluLeuPheGlyProIleThrLeuGluGlnValArgPheProGlyThrGluGlnIle 272
QY 803 ACAGACAGGGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 862
Db 273 ValAsnGluLysGlnLysLeuPheThrSerArgLeuLeuAspValMetAspArgGlyLeu 292
QY 863 GCTCTCTGCGGGCGGGCAGTGGCTCTGGGCGCCAGCGGCTGGGCGACTGCCACATATC 922
Db 293 IleLeuGluValSerGlyHisAlaIleTyrAlaIleArgLeuCysGlnCysLysValTyr 312
QY 923 TGGGCA-----GTGAGCGAGGAGCTCTCCCAACAGCGGCGATGGGCTGATGCG 973
Db 313 TrpSerGlyProCysSerProSerProIleThrProAsnPhe----- 326
QY 974 GAGGTCCCCAAGACAGAAAGAGCGGTGTGTGACCTGGGCGCTTCAATTGTAGATCTG 1033
Db 327 ---IleGluArgGlnLysArgValLysLeuPheCysValGluThrPheLeuSerAspLeu 345
QY 1034 ATTACCTTCACGGAAGA---AGCGAGCGTCCACAGCGTATGCGCTCTGCTCTGCTGTG 1090
Db 346 IleSerHisGlnLysGlyIleThrLysGlnProProTyrGluIleTyrLeuGlyPhe 365
QY 1091 GGGGAGTCTGCGCCCGAGGACCGCGTGGACAGAGCGTGTGTGATGTGTCAAGGTGTG 1150
Db 366 GlyGluGluTrpProAspGlyLysTyrLysGluArgLysLeuIleValGlnIleIle 385
QY 1151 CCACAGTGTCTCAGGCGCTTGTGTAAGTGGCGCGGCTAGGGGCTGCTCTCCCTGGAG 1210
Db 386 ProIleValAlaArgMetIleIleGluMetPheThrGlyAspSerThrArgSerPheAsp 405
QY 1211 AAT---ACTGTGACCTGCATTTCCACAGCCACCTCTCCCTACCTCCGACCCAG 1267
Db 406 SerGlySerIleArgLeuGlnIleSer-----IleProAspIleLysAspAsn 421
QY 1268 TACAAGCGCTACCTGAGGAGCTTGGTGGAGGCGATG---GATTTCCAGGGCGCT 1318
Db 422 IleValSerHisLeuLysHisLeuTyrArgLeuLeuGlnAsnHisGlnGlyPro 439

RESULT 4
S57837
lymphoid-specific interferon regulatory factor - mouse
N;Alternate names: Pip PU.1 interaction partner
C;Species: Mus musculus (house mouse)
C;Date: 28-Oct-1995 #sequence revision 03-Nov-1995 #text_change 09-Jul-2004
C;Accession: S57837; S57838; I49359
R;Matsuyama, T.; Grossman, A.; Mitternuecker, H.W.; Siderovski, D.P.; Kiefer, F.; Kawakami, T.
Nucleic Acids Res. 23, 2127-2136, 1995
A;Title: Molecular cloning of LSIRF, a lymphoid-specific member of the interferon regula
A;Reference number: S57836; MUID:95334364; PMID:7541907
A;Accession: S57837
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-450 <MAT1>
A;Cross-references: UNIPROT:Q64287; EMBL:U20949; NID:G972947; PIDN:AAA75316.1; PID:G9729
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1995
A;Note: only a part of the coding sequence is given
R;Matsuyama, T.; Grossman, A.; Mitternuecker, H.; Siderovski, D.; Kawakami, T.; Kimura, T.
submitted to the EMBL Data Library, June 1994
A;Description: Molecular cloning of LSIRF, a lymphoid-specific member of the interferon
A;Reference number: S57838

```

```

A;Accession: S57838
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-450 <MAT2>
A;Cross-references: EMBL:U11692; NID:G790199; PIDN:AAA75309.1; PID:G790200
R;Eisenbeis, C.F.; Singh, H.; Storb, U.
Genes Dev. 9, 1377-1387, 1995
A;Title: Pip, a novel IRF family member, is a lymphoid-specific, PU.1-dependent transcrip
A;Reference number: I49359; MUID:953117607; PMID:7797077
A;Accession: I49359
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-450 <REG>
A;Cross-references: EMBL:U34307; NID:G976446; PIDN:AAA75283.1; PID:G976447
C;Genetics: 72/3; 135/1; 164/3; 213/1; 249/1; 367/1; 404/3
A;Introns: 72/3; 135/1; 164/3; 213/1; 249/1; 367/1; 404/3
C;Superfamily: lymphoid-specific interferon regulatory factor
C;Keywords: alternative splicing

Alignment Scores:
Pred. No.: 3,23e-17 Length: 450
Score: 388.00 Matches: 124
Percent Similarity: 40.13% Conservative: 55
Best Local Similarity: 27.80% Mismatches: 183
Query Match: 14.38% Indels: 84
DB: 2 Gaps: 17

US-09-975-253-1 (1-1426) x S57837 (1-450)
QY 77 TGGTGTGTTCACAGTCCAGCTGGGCAACTGGAGGGGTGGCTGGCTGGTGAACAGAGC 136
Db 27 TrpLeuIleAspGlnIleAspSerGlyLysTyrProGlyLeuValTrpGluAsnGlu 46
QY 137 CCACGCGCTTCGGATCCCTTGGAGACGCGCTACGCGAGAT---SCACAGCAGGAG 193
Db 47 LysSerValPheArgIleProTrpLysHisAlaGlyLysGlnAspTyrAsnArgGlu 66
QY 194 GATTTCCGGAATCTTCAGGCGCTGGGCGGCGCACTGGTGCATATATTTCCCGGAGGAT 253
Db 67 AspAlaLeuPheLysAlaTrpAlaLeuPheLysGlyLysPheArgGluGlyIleAsp 86
QY 254 AAGCCAGACCTCCACCTGGAGAGGAATTTCCGCTCTGCGCTCAACCGCAAGAGGG 313
Db 87 LysProAspProProThrTrpLysThrArgLeuArgCysAlaLeuAsnLysSerAsnAsp 106
QY 314 TTGCGTTTACAGACAGACCGGAGCAG---GACCCCTCAGACCCACATAAATCTACGAG 370
Db 107 PheGluGluLeuValGluArgSerGlnLeuAspIleSerAspProTyrLysValTyrArg 126
QY 371 TTTGTG-----AACTCAGGAGTTGGGGACTTTTCCGAGCCAGACACCTCT--- 415
Db 127 IleValProGluGlyAlaLysLysGlyAlaLysGlnLeuThrLeuAspAspThrGlnMet 146
QY 416 -----CCGACACCAATGGTGGAGGAGGAGTACTTCTGATATCCCGAGAA 457
Db 147 AlaMetGlyHisProTyrProMetThrAlaProTyrGlySerLeuProAlaGlnGln--- 165
QY 458 GACATTCTGGATCAGTACTGGGTACATGCTGTGGCCCA----- 499
Db 166 -----ValHisAsnTyrMetMetProHisAspArgSerTrpArg 179
QY 500 -----CTCCAGATCCGGGACCCCAAGCGCTGTAGCCCTGAGCCCTGCCCTCAG 553
Db 180 AspTyrAlaProAspGlnSerHisProGluIle-----ProTyrGlnCysProVal 196
QY 554 CCCTCG-----CGAGGCCCGAGCTTGGACAT--- 580
Db 197 ThrPheGlyProArgGlyHisHisTrpGlnGlyProSerCysGluAsnGlyCysGlnVal 216
QY 581 CCCACTCCCTTCCAAACCTGGGGCCCTCTGAGAAC-----CCACTGAAG 625
Db 217 ThrGlyThrPheTyrAlaCysAlaProProGluSerGlnAlaProGlyIleProIleGlu 236

```

```

QY 626 CGGCTGTTGGTCCGGGGNA-----GAGTGGAGTTGAGGTGACAGCC 670
Db      ::::
237 ProSerileArgSerAlaGluAlaLeuAlaLeuSerAspCysArgLeuHisIleCysLeu 256
QY 671 TTCTACCGGGGCCCAAGTCTTCCAGCAGACACCATCTCTGCGCCGAGGSCCTGGCGCTG 730
Db      ::::
257 TyrTyrArgAspileuVallysGluLeuThrThrSerProGluGlyCysArgIle 276
QY 731 GTGGGTCCGAGTGGGAGACAGGAGCGTCCCTGGATGGCCAGTCACTGCCAGACCT 790
Db      ::::
277 SerHisGlyHisThrTyrAspValSerAsnLeuAsp---GlnValLeuPheProTyrPro 295
QY 791 GGCATCTCCCTGACACAGAGGAGTATGATGATGATGATGATGATGATGATGATGAT 850
Db      ::::
296 -----AspAspAsnGlyGlnArgLysAsnIleGluLysLeuSerHisLeu 311
QY 851 GTGGGGGACTGCTCTGCGCGGGCGGCGAGTGGCTCTGGGCCAGCGGCTGGGGCAC 910
Db      ::::
312 GluArgGlyLeuValLeuTyrMetAlaProAspGlyLeuTyrAlalysArgLeuCysGln 331
QY 911 TGCCACACATAGTGGCAGTGGCAGGAGTGGCTCTCCCAACAGCGGGCATGGGCT--- 967
Db      ::::
332 SerArgIleTyrTrp-----AspGlyProLeu 340
QY 968 -----GATGCGAGTCCCAAGCAGCAAGAGGAGCGCTGTTTGAC 1009
Db      ::::
341 AlaLeuCysSerAspArgProAsnLysLeuGluArgAspGlnThrCysLysLeuPheAsp 360
QY 1010 CTGGGGCCCTTCATTGTAGATCTGATTACCTTTCACGAAAGAGCGGCTCACACGC 1069
Db      ::::
361 ThrGlnGlnPheLeuSerGluLeuGlnValPheAlaHisGlyArgProAlaProArg 380
QY 1070 TATGCCCTCTGTTGTTGGGGAGTCAATGGCCCCAGACGACCGTGGACCAAGAG 1129
Db      ::::
381 PheGlnValThrLeuCysPheGlyGluGluPheProAspProGlnArg---GlnArgLys 399
QY 1130 CTCGTGATGTTCAAGTGTGCGCCAGCTGCTCAGGCGCTTGGTAGAATGCCCGGTA 1189
Db      ::::
400 LeuIleThrAlaHisValGluProLeuLeuAlaArgGlnLeuTyrTyrPheAlaGlnGln 419
QY 1190 GGGGTGCTCTCCCTGGAG-----AATACTGTGGACCTG 1225
Db      ::::
420 AsnThrGlyHisPheLeuArgGlyTyrGluLeuProGluHisValThrThrProAspTyr 439
QY 1226 CACATTCCACAGCCAC 1243
Db      ::::
440 HisArgSerLeuArgHis 445

RESULT 5
A45064
interferon consensus sequence binding protein, ICSBP (DNA-binding domain) - human
C:Species: Homo sapiens (man)
C>Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 13-Sep-1998
C:Accession: A45064
J.Weisz, A.; Marx, P.; Sharf, R.; Appella, E.; Driggers, P.H.; Ozato, K.; Levi, B.Z.
R. Biol. Chem. 267, 25589-25596, 1992
A>Title: Human interferon consensus sequence binding protein is a negative regulator of
A:Reference number: A45064; PMID:93094284; PMID:1460054
A:Accession: A45064
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-425 <WEI>
A:Experimental source: lung and blood
A>Note: sequence extracted from NCBI backbone (NCBIP:120312)
C:Superfamily: lymphoid-specific interferon regulatory factor
C:Keywords: DNA binding; transcription regulation

Alignment Scores:
Pred. No.: 7,45e-15 Length: 425
Score: 351.00 Matches: 119
Percent Similarity: 42.17% Conservative: 56
Best Local Similarity: 28.67% Mismatches: 172
Query Match: 13.00% Indels: 68

```

```

DB: 2 Gaps: 15
US-09-975-253-1 (1-1426) x A45064 (1-425)
QY 65 CGGNTCTCCCTGGCTGGTGTGGAGCTGGACCTGGGCACTGGAGGCGTGGCGCTGG 124
Db 9 ArgLeuArgGlnTrpLeuIleGluGlnIleAspSerSerMetTyrProGlyLeuIleTyr 28
QY 125 GTGAACAAGAGCGGACGCGCTTCCGCATCCCTTGAAGCAGCGCTACGCGAGGATGCA 184
Db 29 GluAsnGluGluLysSerMetPheArgIleProTyrLysHisAlaGlyLysGlnAspTyr 48
QY 185 CAGCAGGAG---GATTTCGGAATCTTCCAGGCTGGGCGGAGGCCACTGCTGCATATGTT 241
Db 49 AsnGlnGluValAspAlaSerIlePheLysAlaTyrAlaValPheLysGlyLysPheLys 68
QY 242 CCGGAGGAGGATAAGCCAGCACCTGCCCACTGGAGAGGAATTCGCTCTGCCCTCAAC 301
Db 69 GluGly---AspLysAlaGluProAlaThrTyrLysThrArgLeuArgCysAlaLeuAsn 87
QY 302 CGCAAGAAGGCTTGGCTTTAGCAGAGGACCGAGCAAG---GACCTCACGACCCACAT 358
Db 88 LysSerProAspPheGluValThrAspArgSerGlnLeuAspIleSerGluProTyr 107
QY 359 AAAATCTAGAGTTTGTG-----AATCAGAGTGTGGGACTTT 397
Db 108 LysValTyrArgIleValProGluGluAspGlnLysCysLysLeuGlyValAlaThrAla 127
QY 398 TCCGACCCAGACACCTCTCCGACACCAATGTGGAGGAGTACTTCTGAT-----ACC 451
Db 128 GlyCysValAsnGluValThrGluMetGluCysGlyArgSerGluLeuLeuLeu 147
QY 452 CAGGAAGACATCTGGATGAGTTACTGGGTAAATGCTGTGGCCCACTCCAGATCCG 511
Db 148 LysGluProSerValAspAspTyrMetGlyMetIleLysArgSerProSerPro----- 165
QY 512 GGACCCCAAGCCTGGCTGTAGCCCTGAGCCCTGACCCCTGAGCCCTCGGAGCCCCAGC 571
Db 166 -----ProAspAlaCysArgSerGlnLeu----- 173
QY 572 TTGGACAATCCCACTCCCTTCCCAACCTGGGGCCCTCTGAGAACCCACTGAGCGGCTG 631
Db 174 -----LeuProAspTyrTrpAlaHisGluProSerThrGlyArgArg 187
QY 632 TTGGTCCCGGGGAGAGTGG-----GAGTTCCGAGGTG 664
Db 188 LeuValThrGlyTyrThrThrTyrAspAlaHisHisSerAlaPheSerGlnMetValIle 207
QY 665 ACAGCCTTCTACCGGGGCGCCAGTCTTCCAGCAGACCATCTCTGCGCCGAGGSCCTG 724
Db 208 SerPheTyrTyrGlyGlyLysLeuValGlyGlnAlaThrThrCysProGluGlyCys 227
QY 725 CGGCTGGTGGGTCCGAAGTGGGAGACAGGCGCTGCTGGATGGCCAGTCACTGCA 784
Db 228 ArgLeuSerLeuSerGlnProGly-----LeuProGlyThrLysLeuTyrGlyPro 244
QY 785 GACCTGCGCATGTCCTCTG-----ACAGACAGGGGAGTGTATGATGCTAGCTG 829
Db 245 Glu---GlyLeuGluLeuValArgPheProAlaAspThrIleProSerGluArgGln 263
QY 830 AGGCATGTCTGAGTGTGCTGGTGGGGGACTG-----GCTCTGCGGGCGCGG 880
Db 264 ArgGlnValThrArgLysLeuPheGlyHisLeuGluArgGlyValLeuHisSerArg 283
QY 881 CAGTGGCTCTGGCCCGGCGCTGGGCACTGCCACATACTGGGAGTGGAGGAG 940
Db 284 GlnGlyValPheValLysArgLeuCysGlnGlyArgValPheCys----- 298
QY 941 CTGCTCCCCAACAGCGGCGAT-----GGGCGCTGATGGGAGTGTCCCAAG 985
Db 299 -----SerGlyAsnAlaValValCysLysGlyArgProAsnLysLeuGluArg 314
QY 986 GACAAGGAGGAGGCGTGTGTTGACCTGGGGCCCTTCATTGTAGATCTGATTACCTTACG 1045

```

148	LysGluProSerValAspGluTyrMetGlyMetThrLysArgSerProSerPro	165	
QY	512	GGACCCCAAGCCTGGCTGTAGCCCTGAGCCCTGCCTCAGCCCTCG	559
Db	166	-----ProGluAlaCysArgSerGlnIleLeuProAspTrp	177
QY	560	-----CGAGACCCACGCTGGACAATGCC-----ACTCCCTCCCAACCTGGGGCCC	607
Db	178	TrpValGlnProSerAlaGlyLeuProLeuValThrGlyTyrAlaAlaTyrAspThr	197
QY	608	TCTCAGAACCCACTGAGCGGCTCTGGTGCCTGGGGGAAGAGTGGGAGTTCAGAGTGACA	667
Db	198	HisHisSerAlaPheSerGlnMetVal-----IleSer	208
QY	668	GCCTTCTACCGGGGCGCCCAAGTCTTCCAGCAGACCATCTCCTCGCGAGGGCGCTGGG	727
Db	209	PheTyrIrrGlyGlyLysLeuValGlyGlnAlaThrThrCysLeuGluGlyCysArg	228
QY	728	CTGGTGGGTCCGAAGTGGGAGACAGGACGCTGCTGGA-----TGGCCAGCT	775
Db	229	LeuSerLeuSerGlnProGlyLeuProLysLeuTyrGlyProAspGlyLeuGluProVal	248
QY	776	ACACTGCCAGACCCCTGGCATCTCCCTGACACAGGGGAGTGATGAGCTAGCTGAGGCAT	835
Db	249	CysPheProThrAlaAspThrIleProSerGluArg---GlnArgGlnValThrArgLys	267
QY	836	GTGCTGAGCTGCTGGGTGGGGGACTGCCTCTGCGGGCGGGCAGTGGCTGGGCC	895
Db	268	LeuPheGlyHisLeuGluAluArgGlyValLeuLeuHisSerAsnArgLysGlyValPheVal	287
QY	896	CAGCGGCTGGGGCACTGCCACACATPACTGGGCACTGACGAGGAGCTGCTCCCCAACAGC	955
Db	288	TrpGlnGlyValPheCys-----Ser	298

QY	956	GGGCAT-----GGGCCTGATGGCGAGGTCCCAACGACCAAGGAAGGAGGC	1001
Db	299	GlyAsnAlaValValCysLysGlyArgProAsnLysLeuGluuArgAspGluValValGln	318
QY	1001	GHGTTTGACCTGGGCGCCCTTCATCTAGATCTGATTACTCTCACGGAAGGAAGCGGACGC	1066
Db	319	ValPheAspThrAsnGlnPheIleArgGluLeuGlnInPheTyrAlaThrGlnSerArg	338
QY	1061	TCACCACGCTATGCCTCTGGTCTGTGTGGGGGAGTCTATGCGCCCGACGACGCGCTGG	1122
Db	339	LeuProAspSerArgValValLeuCysPheGlyGluGluPheProAspThrValProLeu	358
QY	1121	ACCAAGAGCCTCGTGTATGGTCAAGGTGTGTGCCACGTCCTCAGGCGCTTGTGTAGAAATG	1188
Db	359	ArgSerLysLeuIleLeuValGlnValGluGlnLeuTyrAlaArgGlnLeuValGluGlu	378
QY	1181	GCC---CGGGTAGGGGGTGCTCCTCCCTG 1207	
Db	379	AlaGlyLysSerCysGlyAlaGlySerLeu 388	

RESULT 7

A45017

transcription factor ISGF3 gamma chain - human

N;Alternate names: alpha-interferon-responsive transcription factor

C;Species: Homo sapiens (man)

C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004

R;Accession: A45017

R;Veale, S.A.; Schindler, C.; Leonard, D.; Fu, X.Y.; Aebersold, R.; Darnell Jr

Mol. Cell. Biol. 12, 3315-3324, 1992

A;Title: Subunit of an alpha-interferon-responsive transcription factor is rel

A;Reference number: A45017; MUID:92334329; PMID:1630447

A;Accession: A45017

A;Status: nucleic acid sequence not shown

A;Molecule type: mRNA

A;Residues: 1-393 <VEA>

A;Cross-references: UNIPROT:Q00978; GB:M87503; NID:G184652; PIDN:AAA58687.1; F

A;Experimental source: HeLa cells

A;Note: sequence extracted from NCBI backbone (NCBIP:108663)

A;Note: part of this sequence was confirmed by protein sequencing

C;Superfamily: lymphoid-specific interferon regulatory factor
C;Keywords: DNA binding; transcription factor

Alignment Scores:

Pred. No.: 1,08e-11 Length: 393
Score: 301.50 Matches: 108
Percent Similarity: 41.46% Conservative: 62
Best Local Similarity: 26.34% Mismatches: 177
Query Match: 11.17% Indels: 63
DB: 2 Gaps: 16

US-09-975-253-1 (1-1426) x A45017 (1-393)

```

QY 38 GGGCGGACCATGGGACCAACCCCAAGCCAGCGTCTGCGCCCTGGCTGGTGTGCTGCGAGCTGGAC 97
Db 4 GlyArgAlaArgCysThrArgLysLeuArgAsn-----TrpValValGluGlnValGlu 21
QY 98 CTGGGCAACTGGAGGCGCTGGCTGGTGAACAAGAGCGCGACGCGCTTCGCGATCCCT 157
Db 22 SerGlyGlnPheProGlyValCysTrpAspAspThrAlaLysThrMetPheArgIlePro 41
QY 158 TGAAGCAGCGCTACGGCAGGAT-----GCACAGCAGGAGATTTCGGAATCTTCCAGGCC 214
Db 42 TrpLysHisAlaGlyLysGlnAspPheArgGluAspGlnAspAlaAlaPhePheYsAla 61
QY 215 TGGGCGGAGCCACTGTCATATGTCCTCCGGAGGGATAGCCACCTGCCAACCCTGG 274
Db 62 TrpAlaIlePheLysGlyLysLysGluGlyLysGluGlyLysGluGlyLysGluGlyLys 80
QY 275 AAGAGGAATTCGCTCTGCTCCCTCAACCGCAAGAGGGTGGTGTAGCAGAGGACCGG 334
Db 81 LysThrArgLeuArgCysAlaLeuAsnLysSerSerGluPheLysGluValProGluArg 100
QY 335 AGCAAG---GACCCTCAGCACCACATAAATCTACAGTTGTGTAACCTCAGAGTGGG 391
Db 101 GlyArgMetAspValAlaGluProTyrLysValTyrGlnLeuLeuProGlyIleVal 120
QY 392 GACTTTTCCCGCCAGACACTCT-----CCGGACACCAATGGTGGAGGCTACTTCT 445
Db 121 Ser---GlyGlnProGlyThrGlnLysValProSerLysArgGlnHisSerValSer 139
QY 446 GATACCCAGAGACATCTCGATGAGTACTGGGTAAACATGTTGGCCCACTC--- 502
Db 140 SerGluArgLysGlu---GluGluAspAlaMetGlnAsnCysThrLeuSerProSerVal 158
QY 503 -----CCAGAT 508
Db 159 LeuGlnAspSerLeuAsnAsnGluGluGlyAlaSerGlyGlyAlaValHisSerAsp 178
QY 509 CCGGACCCCAAGCTGCTAGCCCTGAGCCCTGAGCCCTGCGCTCAGCCCTGGAGCCCC 568
Db 179 IleGlySerSerSerSerSerSerSerProGluPro-----Gln 191
QY 569 AGCTTGGACAAATCCCACTCCCTTCCCAACCTGGGCGCCCTCTGAGAACCCACTGAAGCG 628
Db 192 GluValThrAspThrThrGluAlaProPheGlnGly---AspGlnArgSerLeuGluPhe 210
QY 629 CTGTTGGTCCGGGAGAGAGTGGAGTTCGAGGTGACAGCCCTTACCGGGCGCCGCCA 688
Db 211 LeuLeuProGluProAspTyrSerLeuLeuLeuThrPheIleTyrAsnGlyArgVal 230
QY 689 GTCTTCCAG-----CAGACCATCTCTCCCGGGAGGCTGGCTGGTGGTGGTGGTGG 739
Db 231 ValGlyGluAlaGlnValGlnSerLeuAspCys-----ArgLeuValAlaGlu 246
QY 740 GAAGTGGGAGACAGGACGCTGCTGGATGGCGAGTCACACTGCCAGACCCCTGGCATGTC 799
Db 247 ProSerGlySerGluSerSerMetGlu---GlnValLeuPheProLysPro----- 262
QY 800 CTGACAGACAGGGAGTGAATAGCTACGTGAGCCATGTGCTGAGCTGCTGGTGGGGA 859
Db 263 -----GlyProLeuGluProThrGlnArgLeuLeuSerGlnLeuGluArgGly 278

```

```

QY 860 CTGGCTCTCTGGCGGCGCGGAGTGGCTCTGGGCGCCAGCGGCTGGGCGCACTGCCACACA 919
Db 279 IleLeuValAlaSerAsnProArgLysLeuPheValGlnArgLeuCysProIleProIle 298
QY 920 TACTGGCAGTGAAGCGAGGAGCTGCCCAACAGCGGGCATGGGCTGATGCGGAG--- 976
Db 299 SerTrp-----AsnAlaProGlnAlaProGlyProGly 310
QY 977 -----GTCCCCAAGACAGGAGGAGGCGTGTGGACCTGGGCGCTTCAATTGTA 1027
Db 311 ProHisLeuLeuProSerAsnGluCysValGluLeuPheArgThrAlaTyrPheCysArg 330
QY 1028 GATCTGATTACCTTCCGGAAGACGAGCGCTCACAGCTATGCCCTCTGCTTCTGT 1087
Db 331 AspLeuValArgTyrPheGlnGlyLeuGlyProProLysPheGlnValThrLeuAsn 350
QY 1088 GTGGGGAGTCATGGCCCCCAGGACCGCTGGACCAAGAGGCTCGTGATGGTCAAGTT 1147
Db 351 PheTrpGluGluSerHisLysSerHisThrProGlnAsnLeuIleThrValLysMet 370
QY 1148 GTGCCACCTGCTCAGGCGCTTGGTAGAA 1177
Db 371 GluGlnAlaPheAlaArgTyrLeuLeuGlu 380

RESULT 8
S57836
Lymphoid-specific interferon regulator factor - mouse (fragments)
C;Species: Mus musculus (house mouse)
C;Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 17-Mar-1999
C;Accession: S57836
R;Matsuyama, T.; Grossman, A.; Mittrucker, H.W.; Siderovski, D.P.; Kiefer, F.; Kawakami
Nucleic Acids Res. 23, 2127-2136, 1995
A;Title: Molecular cloning of LSIRF, a lymphoid-specific member of the interferon regula
A;Reference number: S57836; MUID:95334364; PMID:7541907
A;Accession: S57836
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-203 <MAT>
A;Cross-references: EMBL:U11692
C;Superfamily: lymphoid-specific interferon regulatory factor

Alignment Scores:
Pred. No.: 4.18e-10 Length: 203
Score: 276.50 Matches: 77
Percent Similarity: 31.25% Conservative: 33
Best Local Similarity: 21.88% Mismatches: 79
Query Match: 10.24% Indels: 163
DB: 2 Gaps: 6

US-09-975-253-1 (1-1426) x S57836 (1-203)
QY 77 TGGCTGTGTGCGACCTGGGCACTGGAGGCGTGGCTGGTGAACAGAGC 136
Db 5 TrpLeuIleAspGlnIleAspSerGlyLysTyrProGlyLeuValTrpGluAsnGlu 24
QY 137 CGACGCGCTTCGCGATCCCTTGGAGCAGCGCTTACGCGAGGAT---GCACAGCAGGAG 193
Db 25 LysSerValPheArgIleProTyrLysHisAlaGlyLysGlnAspTyrAsnArgGlu 44
QY 194 GATTTCGGAATCTCCAGCGCTGGCGGAGCCACTGTGTCATATGTTCCCGGGAGGAT 253
Db 45 AspAlaAlaPheLysAlaTrpAlaLeuPheLysGlyLysPheArgGluGlyIleAsp 64
QY 254 AAGCCAGACCTGCCAACCTGGAAGAGGAATTTCCGCTCTGCGCTCAACCGCAAGAGGG 313
Db 65 LysProAspProProThrTrpLysThrArgLeuArgCysAlaLeuAsnLysSerAsnAsp 84
QY 314 TTGCGTTTAGCAGAGCCCGGAGCAAG---GACCTCAGCAGCCCAATAAATCTTACGAG 370
Db 85 PheGluGluLeuValGluArgSerGlnLeuAspIleSerAspProTyrLysValTyrArg 104
QY 371 TTTGTGAACCTCAGGAGTGGGGACTTTTCCAGCGCAGACACTCTCCGGAACCAATGGT 430

```


RESULT 9
152998
interferon regulatory factor 1 - human
C:Species: Homo sapiens (man)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C:Accession: I52998
R:Cha, Y.; Sims, S.H.; Romine, M.F.; Kaufmann, M.; Deisseroth, A.B.
DNA Cell Biol. 11, 605-611, 1992
A:Title: Human interferon regulatory factor 1: intron/exon organization.
A:Reference number: I52998; MUID:93000481; PMID:1382447
A:Accession: I52998
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-325 <RES>
A:Cross-references: UNIPROT:P10914; GB:L05072; NID:gi184648; PIDN:AAA36043.1; P
C:Genetics:
A:Cens: GDB:TRF1
A:Cross-references: GDB:127269; OMIM:147575
A:Map position: 5q31.1-5q31.1
A:Introns: 29/3; 63/1; 122/1; 138/3; 182/1; 223/1; 239/3; 285/1
Alignment Scores: 9.13e-08
pred. No.: 325

RESULT 10
JC4592
transcription factor ISGF3 gamma chain - mouse
N:Alternate names: interferon-stimulated gene factor 3 gamma chain
C:Species: Mus musculus (house mouse)
C:date: 10-Apr-1995 #sequence_revision 24-Oct-1998 #text_change 09-Jul-2004
C:Accession: JC4592; S71599
J:Biochem. 119, 231-234, 1996
J:Suhtara, W.; Yoneyama, M.; Yonekawa, H.; Fujita, T.
A:Title: Structure of mouse interferon stimulated gene factor 3 gamma (ISGF3 gamma)
A:Reference number: JC4592; MUID:97037063; PMID:8882710
A:Accession: JC4592
A:Molecule type: mRNA
A:Residues: 1-339 <SU>
A:Cross-references: UNIPROT:061179; EMBL:U51592; NID:G1263309; PIDN:AAC52494.1
A:Experimental source: L929 cells
R:Kawakami, T.; Matsumoto, M.; Sato, M.; Harada, H.; Taniguchi, T.; Kitagawa, I.
FEBS Lett. 358, 225-229, 1995
A:Title: Possible involvement of the transcription factor ISGF3gamma in virus-
A:Reference number: S71599; MUID:95145714; PMID:7843405
A:Accession: S71599

US-09-0975-253-1 (1-1426) x B31595 (1-325)

QY 476 GTAACATCATCAGAAATGCTTCTCGGTATCAGAGTACTGCCTCCACCATGGTGTCCG 417
 Db 1342 AlapAlaProAlaProAlaProAlaProSerSerGlyGlyTyrSer 1355
 QY 416 GAGAGGTGTCTGGCTGGGAAAGTCCCAACTCTCTGAGTTCACAAACTCGTAGATTAT 357
 Db 1356 -----GlyGlySerSer----- 1359
 QY 356 GTGGGTGTGAGGTCTCTCTCGGTCTCTCTGCTGCTTAAACGAAACCCCTTTCTGGGTGA 297
 Db 1360 -----GlyGlySerAlaAlaGly----- 1365
 QY 296 GGGCAGAGCGGAAATCTCTCCAGGTGGCAGGTCTGCTTATCCCTCCCGGAACAT 237
 Db 1366 GlyGlyGlySerSerGlyGlyTyrThrGlyGlySerAlaAlaPro----- 1381
 QY 236 ATGCACAGTGGCTCGGCCCGCCAGCCCTGGAAGATTCGAAATCTCTCTGCTGTGCACTCT 177
 Db 1382 -----ProProProProProProProProProProProProAlaProAlaPro 1397
 QY 176 GCGTAGCGCTCTCCCAAGGATG-----GGAAGCGGTGC--- 138
 Db 1398 AlaProAlaProSerSerGlyGlyTyrSerGlyGlySerSerGlyGlySerAlaAlaGly 1417
 QY 137 -----GGCTCTGTTCACCCAGGCCACGCCCT 111
 Db 1418 GlyGlyGlySerSerGlyGlyTyrSerGlyGlySerAlaAlaPro---ProProPro 1436
 QY 110 CGAGTTGCCCGAGTCCAGTCGCAGCACCGAGCGGAGGAGCCGCTGGCTTTGGGGTTC 51
 Db 1437 ProProAlaProAlaProAlaProAlaProSerSerGlyGlyTyrSerGlyGlySer 1456
 QY 50 CCATGTCGCGCTACGATCGAAGTTCGGGGCGTGGCGGAGCTGG 5
 Db 1456 rGlyGlySerAlaAlaGlyGlyGlyGlyGlySerSerGlyGlyTyr 1471

RESULT 13
 A53340
 N:Interferon regulatory factor 2 - human
 C:Species: Homo sapiens (man)
 C:Date: 03-May-1994 #sequence revision 03-May-1994 #text_change 09-Jul-2004
 C:Accession: A53340; S06894; A32828
 R:Cha, Y.; Deisseroth, A.B.
 J. Biol. Chem. 269, 5279-5287, 1994
 A:Title: Human interferon regulatory factor 2 gene. Intron-exon organization and function
 A:Reference number: A53340; MUID:94148994; PMID:8106512
 A:Accession: A53340
 A:Status: not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-349 <CHA>
 A:Cross-references: UNIPROT:P14316; GB:I24442
 R:Itob, S.; Harada, H.; Fujita, T.; Mimura, T.; Taniguchi, T.
 Nucleic Acids Res. 17, 8372, 1989
 A:Title: Sequence of a cDNA coding for human IRF-2.
 A:Reference number: S06894; MUID:90045964; PMID:2813069
 A:Accession: S06894
 A:Molecule type: mRNA
 A:Residues: 1-57, 'R', 59-349 <ITO>
 A:Cross-references: EMBL:X15949; NID:g33966; PIDN:CAA34073.1; PID:g33967
 R:Harada, H.; Fujita, T.; Miyamoto, M.; Kimura, Y.; Maruyama, M.; Furia, A.; Miyata, T.;
 Cell 58, 729-739, 1989
 A:Title: Structurally similar but functionally distinct factors, IRF-1 and IRF-2, bind to
 A:Reference number: A32828; MUID:89354547; PMID:2475256
 A:Accession: A32828
 A:Molecule type: mRNA
 A:Residues: 1-68, 'I', 70-96, 'R', 98-130, 'ER', 133-152, 'GF', 155-163, 'A', 165-188, 'D', 190-210,
 8, 'T', 310-313, 'PAPY', 318-319, 'TP', 322-349 <HAP>
 A:Cross-references: GB:J03168; NID:g198455; PIDN:AAA39333.1; PID:g293676
 A:Experimental source: clones 2 and 5
 C:Genetics:
 A:Gene: GDB:IRF2
 A:Cross-references: GDB:127270; OMIM:147576

A:Map position: 4q35.1-4q35.1
 C:Keywords: DNA binding; transcription regulation

Alignment Scores:
 Pred. No.: 2, 01e-06 Length: 349
 Score: 219.00 Matches: 79
 Percent Similarity: 39.68% Conservative: 44
 Best Local Similarity: 25.48% Mismatches: 111
 Query Match: 8.11% Indels: 76
 DB: 2 Gaps: 14

US-09-975-253-1 (1-1426) x A53340 (1-349)

QY 47 ATGGGAACCCCAAGAACCCAGCGNTCTCCCTGCTGTGTCTGCAGCTGGACCTGGGGCAA 106
 Db 1 MetProValGluArgMetArgMetArgProTrpLeuGluGluGlnIleAsnSerAsnThr 20
 QY 107 CTGGAGGGGTGGCTGGGTGAACAAGACCGCAGCGCTTCCGATCCCTTGGAAACAC 166
 Db 21 IleProGlyLeuLysTrpLeuAsnLysGluLysLysIlePheGlnIleProTrpMetHis 40
 QY 167 GGCTTACGGCAGGATGCACAG---CAGGAGGATTTCGGAATCTCCAGGCTGGGGCAG 223
 Db 41 AlaAlaArgHisGlyTrpAspValGluLysAspAlaProLeuPheArgAsnTrpAlaIle 60
 QY 224 GCCACTGTGTCATATGTTCCCGGGGAGTAAGCCAGACCTGCCAACCTGGGAAGAGAA 283
 Db 61 HisThrGlyLysHisGlnProGlyValAspLysProAspProLysThrTrpLysAlaAsn 80
 QY 284 TTCCGCTCTGCCCTCAACCGCAAGAGGGTTCGGTTTAGCAGAGACCGGAGC---AAG 340
 Db 81 PheArgCysAlaMetAsnSerLeuProAspIleGluGluValLysAspLysSerIleLys 100
 QY 341 GACCTTCACGACCCACATAAAATCTACAGATT----- 373
 Db 101 LysGlyAsnAlaPheArgValLysArgMetLeuProLeuSerGluArgProSerLys 120
 QY 373 ----- 373
 Db 121 LysGlyLysLysProLysThrGluLysGluAspLysValLysHisIleLysGlnGluPro 140
 QY 374 -----GTGAACCTCAGAGTTGGGACTTTTCCAGCCAGAC----- 409
 Db 141 ValGluSerSerLeuGlyLeuSerAsnGlyValSerAspLeuSer---ProGluTrpAla 159
 QY 410 -----ACCTCTCCCGACACCAATGGTGAGGACGACTCTCTGAT-----ACCCAG 454
 Db 160 ValLeuThrSerThrIleLysAsnGluValAspSerThrValAsnIleIleValValGly 179
 QY 455 GAAGACATTTGGATGAGTTACTGGGTAACTGGTGTGGTGGCCCTCCAGATCCGGGA 514
 Db 180 GlnSerHisLeuAspSerAsnIleGluAsnGlnGluIleVal-----ThrAsn 195
 QY 515 CCCCCAAGCCTGGCTAGCCCTGAGCCCTGAGCCCTGAGCCCTGAGCCCTGAGCCCTGAG 574
 Db 196 ProProAspIleCysGlnValValGlu-----ValThrThrGluSerAsp 210
 QY 575 GACAAATCCCTCTCCCAAACTGGGGCTCTGAG---AACCCACTGAAGCGGCTG 631
 Db 211 GluGlnProValSerMetSerGluLeuLysTrpLeuGlnIleSerProValSerSerTyr 230
 QY 632 TTG-----GTCCGGGGGAGAGTGGAGTTCGAGTTCGAGTTCAGAGCC 670
 Db 231 AlaGluSerGluThrThrAspSerValProSerAspGluGluSerAlaGluGlyArgPro 250
 QY 671 TTCTACCGGGCGGCAAGTCTTCCAGCAGACCATCTCTCTGCCCGGAGGGCTCGCGCTG 730
 Db 251 HisTrpArgLysArgAsnIle-----GluGlyLysGlnTyr 262
 QY 731 GTGGGTCCGAAGTGGGAGACAGG-----ACGCTGCTGGATGCGCA-----GTC 775
 Db 263 Leu---SerAsnMetGlyThrArgLysSerTyrLeuLeuProGlyMetAlaSerPheVal 281

QY 776 ACACCTGACACCTGGCATGTCCCTGACA 805
 Db 282 ThrSerAsnLysProAspLeuGlnValThr 291

RESULT 14
 S13301
 collagen alpha 1(X) chain precursor - bovine
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 21-Nov-1993 #sequence_revision 23-Feb-1996 #text_change 09-Jul-2004
 C:Accession: S13301
 R:Thomas, J. T.; Kwan, A. P. L.; Grant, M. E.; Boot-Handford, R. P.
 Biochem. J. 273, 141-148, 1991
 A:Title: Isolation of cDNAs encoding the complete sequence of bovine type X collagen. EV
 A:Reference number: S13301; MUID:91113131; PMID:1703407
 A:Accession: S13301
 A:Molecule type: mRNA
 A:Residues: 1-674 <THO>
 A:Cross-References: UNIPROT:P23206; EMBL:X53556; NID:g263; PIDN:CAA37624.1; PID:g264
 C:Genetics:
 A:Gene: COL10A1
 C:Superfamily: collagen alpha 1(VIII) chain; complement C1q carboxyl-terminal homology
 C:Keywords: coiled coil; extracellular matrix; glycoprotein; homotrimer
 F:1-18/Domain: signal sequence #status predicted <SIG>
 F:19-674/Product: collagen alpha 1(X) chain #status predicted <MAT>
 P:547-673/Domain: complement C1q carboxyl-terminal homology <CIQ>

Alignment Scores: 674
 Pred. No.: 3.7e-06 Length: 139
 Score: 215.00 Matches: 33
 Percent Similarity: 34.06% Conservative: 184
 Best Local Similarity: 27.52% Mismatches: 149
 Query Match: 8.03% Gaps: 31
 DB: 2

US-09-975-253-1 (1-1426) x S13301 (1-674)

QY 1340 GAGCGAGGGCTCAGCTCTCCAGGGCCCTGGAAATCCA-----TGCCCTCCACCAAGT 1287
 Db 58 GluGlnGlyLeuProGlyProGlyProAlaGlyProArgGlyHisProGlyProSer 77

QY 1286 ---CCTCAGTAGGCTTGTACTGCTCGGAGTGGAGGAGGCTGTGTGAAA 1230
 Db 78 GlyProGlyLysProGlyThrGly----- 86

QY 1229 TGTGACAGTCCACAGTATTCTCCAGGAGGAGGACCCCTTACCCGGCCCATTTCTACCA 1170
 Db 87 SerProGlyProGlnGlyGlnProGly-----LeuProGlyPro-----Pro 100

QY 1169 AGGCCCTGAGGCA-----CG 1155
 Db 101 -GlyProSerAlaThrGlyLysProGlyLeuProGlyLeuProGlyLysGlnGlyGluAr 120

QY 1154 TGGGCACACCTTCCACCATCAGGAGCTTGTCCAGGGTGTCTGGGGGCA---TG 1098
 Db 120 gGlyLeuAsn-----GlyProLysGlyAspIleGlyProAlaGly 133

QY 1097 ACTCCCCACACAGAA-----CCAGAGGGCATACGCTGTGTGAGCGTCCGCTTC 1050
 Db 133 yLeuProGlyProArgGlyProProGlyProGlyLeuProGlyProAlaGlyIleSe 153

QY 1049 CTTC-----CGTGAAGTATACATCTACATGAAGGGCCCGCCAGTCAACACGCTC 996
 Db 153 rValProGlyLysProGlyProGln-----GlyProThrGlyGluProGlyPr 169

QY 995 CTTCCTTGTCTTGGGACCTCCCATCAGGCCCATGCCGCTGTGTGGGAGCAGCTC-- 938
 Db 169 oArgGlyPheProGlyGlyLysGlyThrSerGlyValProGlyLeuAsnGlyGlnLysGl 189

QY 937 ----CTCGCTCACTGCCAGTATGT-----GTGGCAGTGGCCCGCCAGCTGGG 894
 Db 189 yGluMetGlyHisCysThrProCysArgProGlyGluArgGlyLeuProGlyProGlnGl 209

QY 893 CCCAGAGCCTAGTGGCCCGGGCCGCCA-----GAGAGCCAGTCCCCCAC 852

Db 209 yProThrGlyProGlyProGlyValGlyLysArgGlyGluAsnGlyLeuProGl 229
 QY 851 CCAGGACGCTCAGCAGCATGCTCAGTAGCTCATCACTCCCTGCTCTCTCAGGACATGC 792
 Db 229 yGln-----ProGlyLeuLysGlyAsp-----GlnGly-ValP 240

QY 791 CAGGCTCTGGCAGTGTGACTGGCCATCAGGAGCGGTCTCTCTCCCACTTCGGACCCCA 732
 Db 240 roGlyGluArgGlyAlaAlaGlyProSerGly-----ProGlnGlyProProG 256

QY 731 CCAGCCGAGGCGCT-----CGGCGAGAGAGTGTCTGTCTGCTGAGAGACTT 587
 Db 256 yGluGln-GlyProGluGlyLeuGlyLysProGlyAlaProGlyLysProGlyGlnPro 275

QY 686 GGGCGCCCGGTAGAGGCTCAGCTCAGACTCCCACTCTTCCCGCCGACCAACAGCC 627
 Db 276 GlyIleProGlyMetLysGlyGlnProGlyAlaProGlyThrAlaGlyLeuProGlyAla 295

QY 626 GCTTCAGTGGGTCTCAGAGGCGCCAGGTTTG---GGAAGGAGTGGGATTGTCCAAGC 570
 Db 296 ProGlyPheGlyLysProGlyLeuProGlyLeuLysGlyGlnArgGlyProValGlyLeu 315

QY 569 TGGGGCTCCGAGGGCTGAGGCGAGGGCTCAGGGCTA-----CAGCCA 525
 Db 316 ProGlySerProGlyAlaLysGlyGlu---GlnGlyProAlaGlyHisProGlyGluAla 334

QY 524 GCTTGGGGGTCCCGATCTGGAGTGGGCGGCAACACCATGTTACCACTCACTATCCA 465
 Db 335 GlyLeuProGlyProSerGlyAsnMetGlyPro-----GlnGlyPro 348

QY 464 GAATGCTTCTCGGTATCAGAAGTACTGCCTCCACATTGTTGTCGGAGAGGTGTCTG 405
 Db 349 LysGlyIleProGly----- 353

QY 404 GCTGGGAAAGTCCCAACTCTGAGTTCAAACTCTAGATTATTTATGGGTCTGTAG 345
 Db 354 -----AsnProGlyLeuProGlyProLysGlyGlu-----MetGlyProVal 367

QY 344 GTCCTTGTCTCCGCTCTCTGCTAAACGCAACCTCTTTGCGGTTGAGGGCAGAGCGGA 285
 Db 368 GlyProAlaGlyAsnPro-----GlyAlaLysGly 377

QY 284 ---AATTCCTCTCCAGTTTGGCAGTCTG---GTTATCCCTCCCGGGAACATATGCAC 231
 Db 378 GluArgGlySerSerGlyLeuAspGlyLysProGlyTyProGlyGluProGlyLeuAsn 397

QY 230 -----CAGTGGCTCGGCCCGCCAGGCTGGAAGATCCGAAT 195
 Db 398 GlyProLysGlyAsnProGlyLeuProGlyProLysGlyAspProGlyIleAlaGlySer 417

QY 194 CTCTCTGTGCTATCTCGGTAGCGGTCTTCCAGGAGTCCGGAAGCGCGTGGCGG 135
 Db 418 ProGlyLeuProGlyProValGlyProAlaGlyAlaLysGlyValProGlyHisAsnGly 437

QY 134 TCTTGTTCACCCAGG-----CCAGCGCTCCAGTTCGCCAGTCCAGCTGCCACACCA 81
 Db 438 GluAlaGlyProArgGlyValProGlyIleProGlyThrArgGlyProIleGlyProPro 457

QY 80 GCCAGGCGAGGAGCGGTGGTTCCTGCTCCAGTGGTCCGCTACGATGGAAGTCCGG 21
 Db 458 Gly-----IleProGlyPhePro-----GlySerLysGlyAspValGly 470

QY 20 GCGTGGCGG 12
 Db 471 ThrProGly 473

RESULT 15
 S15435
 collagen alpha 1(VIII) chain precursor - human
 C:Species: Homo sapiens (man)
 C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
 C:Accession: S15435

Search completed: October 28, 2004, 15:43:01
Job time : 66.5 secs

This Page Blank (copy)

LENGTH: 427 amino acids	
TYPE: amino acid	
TOPOLOGY: linear	
MOLECULE TYPE: protein	
US-08-705-771-16	
Alignment Scores:	
Pred. No.:	3.82e-178
Score:	2326.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	86.18%
DB:	3
US-09-975-253-1 (1-1426) x US-08-705-771-16 (1-427)	
QY	47 ATGGGACCCCAAGCAGGNTCCCTGCTGGTGTGGAGCTGGACCTGGGGCAA 106
Db	1 MetGlyThrProLysProArg***LeuProTrpLeuValSerGlnLeuAspLeuGlyGln 20
QY	107 CTGGAGGGCGTGGCTGGGTGAACAGAGCGCAGCGCTTCGCAATCCCTTGGAAAGCAC 166
Db	21 LeuGluGlyValAlaTrpValAsnLysSerArgThrArgPheArgIleProTrpLysHis 40
QY	167 GGCCTACGGCAGGATGACAGCAGAGAGATTTCGGAATCTTCAGGCTGGGCGGAGGCC 226
Db	41 GlyLeuArgGlnAspAlaGlnGlnGluAspPheGlyIlePheGlnAlaTrpAlaGluAla 60
QY	227 ACTGGTGATATCTCCGGGAGGGGATAAGCCAGAGCTGCCAACCCTGGAGAGGAATTTC 286
Db	61 ThrGlyAlaTrpValProGlyArgAspLysProAspLeuProThrTrpLysArgAsnPhe 80
QY	287 CGCTCTGCCCTCAACCGCAAGAGGGTTCGTTTAGCAGAGCAGGACCAAGGACCTTG 346
Db	81 ArgSerAlaLeuAsnArgLysGluGlyLeuArgLeuAlaGluAspArgSerLysAspPro 100
QY	347 CACGACCCCATAAATCTACGAGTTGTGAATCTCAGGAGTTGGGACTTTTCCAGCCA 406
Db	101 HisAspProHisLysIleTyrGluPheValAsnSerGlyValGlyAspPheSerGlnPro 120
QY	407 GACACCTCTCGGACCAACTGGTGGAGGAGTACTTCTGATACCCAGGAAGACATCTGT 466
Db	121 AspThrSerProAspThrAsnGlyGlyGlySerThrSerAspThrGlnGluAspIleLeu 140
QY	467 GATGAGTTACTGGGTAACATGTTGTGGCCCACTCCACAGATCCGGACCCCAAGCTGT 526
Db	141 AspGluLeuLeuGlyAsnMetValLeuAlaProLeuProAspProGlyProProSerLeu 160
QY	527 GCTGTAGCCCTGAGCCCTGCCCTCAGCCCTGGGAGCCCGCCAGCTTGGACAATCCCACT 586
Db	161 AlaValAlaProGluProCysProGlnProLeuArgSerProSerLeuAspAsnProThr 180
QY	587 CCCTTCCCAACCTCGGGCCCTCTGAGAACCACTGAAGCGGCTGTGGTCCCGGGGAA 646
Db	181 ProPheProAsnLeuGlyProSerGluAsnProLeuLysArgLeuLeuValProGlyGlu 200
QY	647 GAGTGGGAGTTTCGAGGTGACAGCTTCTACCGGGCGGCCCAAGTCTTCCAGCAGACCATC 706
Db	201 GluTrpGluPheGluValThrAlaPheTyrArgGlyArgGlnValPheGlnGlnThrIle 220
QY	707 TCCTGCGCGGAGGGCTCGGCTGTGGGTCCGAGTGGGAGGAGCAGGACGCTGCTGGA 766
Db	221 SerCysProGluGlyLeuArgLeuValGlySerGluValGlyAspArgThrLeuProGly 240
QY	767 TGGCAGTCACACTCCAGACCCCTGGCATGTCCCTGACAGACAGGGAGTGTATGAGCTAC 826
Db	241 TrpProValThrLeuProAspProGlyMetSerLeuThrAspArgGlyValMetSerTyr 260
QY	827 GTGAGGCATGTCTGAGCTGCTGGGTGGGAGCTGGCTCTCTGCGGCGGGCGGAGTGG 886
Db	261 ValArgHisValLeuSerCysLeuGlyGlyGlyLeuAlaLeuTrpArgAlaGlyGlnTrp 280
QY	887 CTCTGGCCCGGCGGTGGGGGACTGCCACATACTGGGCACTGGGAGTGAGTGTCTC 946

Db	281 LeuTrpAlaGlnArgLeuGlyHisCysHisThrTyrTrpAlaValSerGluGlnLeuLeu 300
QY	947 CCCAACACGGCGCATGGCCCTGATGGCGAGGTCCCAAGGCAAGCAAGGAGGCGCTGTTT 1006
Db	301 ProAsnSerGlyHisGlyProAspGlyGluValProLysAspLysGluGlyValPhe 320
QY	1007 GACCTGGGGCCCTTCATTGTAGATCTGATTACCTTCACGGAAGAGCGGACGTCACCA 1066
Db	321 AspLeuGlyProPheIleValAspLeuIleThrPheThrGluGlySerGlyArgSerPro 340
QY	1067 CGCTATGCCCTCTGTTCTGTGTGGGAGTGTATGCCGCCAGCAGCAGCCCTGGACCAAG 1126
Db	341 ArgTyrAlaLeuTrpPheCysValGlyGluSerTrpProGlnAspGlnProTrpThrLys 360
QY	1127 AGGCTCGTGTGTCAAGTTGTGCTCCAGCTCCCTCAGGCGCTTGGTAGAATGGCCCGG 1186
Db	361 ArgLeuValMetValLysValValProThrCysLeuArgAlaLeuValGluMetAlaArg 380
QY	1187 GTAGGGGGTGCCTCTCCCTGGAGAACTACTGTGGACCTGCACATTTCCAAACGACCCCA 1246
Db	381 ValGlyGlyAlaSerSerLeuGluAsnThrValAspLeuHisIleSerAsnSerHisPro 400
QY	1247 CTCTCCCTCACCCTCCAGCCAGTACAGGCTTACCTGCAGGACTTGTGGAGGCGATGGAT 1306
Db	401 LeuSerLeuThrSerAspGlnTyrLysAlaTyrLeuGlnAspLeuValGluGlyMetAsp 420
QY	1307 TTCCAGGGCCCTGGGAGGAC 1327
Db	421 PheGlnGlyProGlyGluSer 427

RESULT 2

US-09-417-540-16

; Sequence 16, Application US/09417540

; Patent No. 6639052

; GENERAL INFORMATION:

; APPLICANT: Paul Moore, Reiner Gentz, Hongjin Ji,

; Jian Ni and Jing-Shan Hu

; TITLE OF INVENTION: Human Genes, Sequences and

; Expression Products

; NUMBER OF SEQUENCES: 22

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,

; CECCHI, STEWART & OLSTEIN

; STREET: 6 BECKER FARM ROAD

; CITY: ROSELAND

; STATE: NEW JERSEY

; COUNTRY: USA

; ZIP: 07068

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5 INCH DISKETTE

; COMPUTER: IBM PS/2

; OPERATING SYSTEM: MS-DOS

; SOFTWARE: WORD PERFECT 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/417,540

; FILING DATE: 14-Oct-1999

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/705,771

; FILING DATE: August 30, 1996

; ATTORNEY/AGENT INFORMATION:

; NAME: MULLINS, J.G.

; REGISTRATION NUMBER: 33,073

; REFERENCE/DOCKET NUMBER: 325800-346 (PF196)

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 973-994-1700

; TELEFAX: 973-994-1744

; INFORMATION FOR SEQ ID NO: 16:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 427 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

Percent Similarity:	98.13%	Conservative:	0
Best Local Similarity:	98.13%	Mismatches:	1
Query Match:	20.88%	Indels:	1
DB:	3	Gaps:	1
US-09-975-253-1 (1-1426) x US-08-654-482-7 (1-107)			
QY	65	CGGTCCTCGCTGGCTGGTGTGGAGCTGGGCGCACTGGAGGGCGTGGCTGG	124
Db	1	ArgileLeuProTyrValSerGlnLeuAspLeuGlnLeuGluGlyValAlaTyr	20
QY	125	GTGAACAGAGCGGACGCGCTTCGCGATCCCTGGAGACCGGCTACGGCAGGATGCA	184
Db	21	ValAsnLysSerArgThrArgPheArgGileProTyrLysHisGlyLeuArgGlnAspAla	40
QY	185	CAGCAGGAGGATTCGGATCTTCAGGCTGGGCGGAGCCACTGGTGCATATGTCTCC	244
Db	41	GlnGlnGluAspPheGlyIlePheGlnAlaTyrAlaGluAlaThrGlyAlaTyrValPro	60
QY	245	GGGAGGATAAGCCAGACCTGCCAACCTCGAAGAGGAATTCGCG	301
Db	61	GlyArgAspLysProAspLeuProThrTyrLysArgAsnPheArgSerSerAlaLeuAsn	80
QY	302	CGCAAGAGAGGTTGGTGTAGCAGAGGACCGGACGACCTCAGCAGCCACATAAA	361
Db	81	ArgLysGluGlyLeuArgLeuAlaGluAspArgSerLysAspProHisAspProHisLys	100
QY	362	ATCTACGATTTGTGAACCTCA 382	
Db	101	IleTyrGluPheValAsnSer 107	
RESULT 4			
US-09-513-999C-5034			
; Sequence 5034, Application US/09513999C			
; Patent No. 6783961			
; GENERAL INFORMATION:			
; APPLICANT: Dumas Milne Edwards, J.B.			
; APPLICANT: Duclert, A.			
; APPLICANT: Giordano, J.Y.			
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.			
; Patent No. 6783961			
; FILE REFERENCE: 59. US2. REG			
; CURRENT APPLICATION NUMBER: US/09/513.999C			
; CURRENT FILING DATE: 2000-02-24			
; PRIOR APPLICATION NUMBER: US 60/122,487			
; PRIOR FILING DATE: 1999-02-26			
; NUMBER OF SEQ ID NOS: 3681			
; SOFTWARE: Patent.pm			
; SEQ ID NO 5034			
; LENGTH: 102			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: UNSURE			
; LOCATION: 99			
; OTHER INFORMATION: Xaa=Asp or Glu			
US-09-513-999C-5034			
Alignment Scores:			
Pred. No.:	9.2e-36	Length:	102
Score:	546.00	Matches:	100
Percent Similarity:	98.04%	Conservative:	0
Best Local Similarity:	98.04%	Mismatches:	2
Query Match:	20.23%	Indels:	0
DB:	4	Gaps:	0
US-09-975-253-1 (1-1426) x US-09-513-999C-5034 (1-102)			
QY	47	ATGGGAACCCCAAGCAGCGTTCCTGGCTGGTGTGGAGCTGGAGCTGGGCGAA	106
Db	1	MetGlyThrProLysProArgileLeuProTyrValSerGlnLeuAspLeuGlyGln	20
QY	107	CTGGAGGCGTGGCTGGTGAACAGCCGACCGCTTCGCGATCCCTTGGAGGCAC	166
US-09-975-253-1 (1-1426) x US-08-654-482-12 (1-100)			
QY	896	CAGCGCTGGGCACTGCCACATACCTGGGCGAGTGAGCGAGCTGCCCAACAGC	955
Db	1	GlnArgLeuGlyHisCysHisThrTyrTrpAlaValSerGluGluLeuProAsnSer	20
US-08-654-482-12			
Alignment Scores:			
Pred. No.:	4.81e-35	Length:	100
Score:	537.00	Matches:	98
Percent Similarity:	98.00%	Conservative:	0
Best Local Similarity:	98.00%	Mismatches:	2
Query Match:	19.90%	Indels:	0
DB:	3	Gaps:	0
US-08-654-482-12			
Computer Readable Form:			
MEDIUM TYPE: Floppy disk			
COMPUTER: IBM PC compatible			
OPERATING SYSTEM: PC-DOS/MS-DOS			
SOFTWARE: PatentIn Release #1.0, Version #1.30			
CURRENT APPLICATION DATA:			
APPLICATION NUMBER: US/08/654,482			
FILING DATE: 28-MAY-1996			
CLASSIFICATION: 435			
ATTORNEY/AGENT INFORMATION:			
NAME: White, John P.			
REGISTRATION NUMBER: 28,678			
REFERENCE/DOCKET NUMBER: 50995			
TELECOMMUNICATION INFORMATION:			
TELEPHONE: (212) 278-0400			
TELEFAX: (212) 391-0525			
INFORMATION FOR SEQ ID NO: 12:			
SEQUENCE CHARACTERISTICS:			
LENGTH: 100 amino acids			
TYPE: amino acid			
STRANDEDNESS: single			
TOPOLOGY: linear			
MOLECULE TYPE: peptide			
US-08-654-482-12			
GENERAL INFORMATION:			
APPLICANT: Dalla-Favera, Riccardo			
TITLE OF INVENTION: IDENTIFICATION OF GENES ALTERED IN			
TITLE OF INVENTION: MULTIPLE MYELOMA			
NUMBER OF SEQUENCES: 17			
CORRESPONDENCE ADDRESS:			
ADDRESSEE: Cooper & Dunham LLP			
STREET: 1185 Avenue of the Americas			
CITY: New York			
STATE: New York			
COUNTRY: U.S.A.			
ZIP: 10036			
US-08-654-482-12			
; Sequence 12, Application US/08654482			
; Patent No. 6245562			
; GENERAL INFORMATION:			
; APPLICANT: Dalla-Favera, Riccardo			
; TITLE OF INVENTION: IDENTIFICATION OF GENES ALTERED IN			
; TITLE OF INVENTION: MULTIPLE MYELOMA			
; NUMBER OF SEQUENCES: 17			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Cooper & Dunham LLP			
; STREET: 1185 Avenue of the Americas			
; CITY: New York			
; STATE: New York			
; COUNTRY: U.S.A.			
; ZIP: 10036			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: PatentIn Release #1.0, Version #1.30			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/08/654,482			
; FILING DATE: 28-MAY-1996			
; CLASSIFICATION: 435			
; ATTORNEY/AGENT INFORMATION:			
; NAME: White, John P.			
; REGISTRATION NUMBER: 28,678			
; REFERENCE/DOCKET NUMBER: 50995			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: (212) 278-0400			
; TELEFAX: (212) 391-0525			
; INFORMATION FOR SEQ ID NO: 12:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 100 amino acids			
; TYPE: amino acid			
; STRANDEDNESS: single			
; TOPOLOGY: linear			
; MOLECULE TYPE: peptide			
US-08-654-482-12			
Alignment Scores:			
Pred. No.:	4.81e-35	Length:	100
Score:	537.00	Matches:	98
Percent Similarity:	98.00%	Conservative:	0
Best Local Similarity:	98.00%	Mismatches:	2
Query Match:	19.90%	Indels:	0
DB:	3	Gaps:	0

QY 956 GGGCATGGGCTGATGCGAGGTCCCAAGGACAAGAGGAGCGGTGTTGACCTGGGG 1015
Db 21 GlyHisGlyProAspGlyGluValProLysAspLysGluGlyValPheAspLeuGly 40
QY 1016 CCTTCATTGTAGATCTGATTACCTTCACGGAAGGAGCGGAGCTCACCACGCTATGCC 1075
Db 41 ProPheIleValAspLeuIleThrPheThrGluGlySerGlyArgSerProArgTyrAla 60
QY 1076 CTCTGTTTGTGTGGGAGTATGCGCCCGAGGACGCGGTGGACCAAGAGCTCGTG 1135
Db 61 TrpLeuPheCysValGlyGluSerTrpProGlnAspGlnProTrpThrLysArgLeuVal 80
QY 1136 ATGGTCAAGGTTGTGCCCGAGTCCCTCAGGCGCTTGGTGAATGCGCGGTAGGGGT 1195
Db 81 MetValLysValValProThrCysLeuArgAlaLeuValGluMetAlaArgValGlyGly 100

RESULT 6

US-08-999-774A-10
; Sequence 10, Application US/08999774A
; Patent No. 6274312
; GENERAL INFORMATION:
; APPLICANT: Gish, Kurt C.
; APPLICANT: Seghezzi, Wolfgang
; APPLICANT: Shanahan, Frances
; APPLICANT: Lees, Emma M.
; APPLICANT: McClanahan, Terrill K.
; TITLE OF INVENTION: Intracellular Regulatory Molecules;
; NUMBER OF INVENTION: Related Reagents
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DNAX Research Institute
; STREET: 901 California Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/999,774A
; FILING DATE: 10-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/032,818
; FILING DATE: 11-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ching, Edwin P.
; REGISTRATION NUMBER: 34,090
; REFERENCE/DOCKET NUMBER: DX0646
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650)852-9196
; TELEFAX: (650)496-1200
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 503 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-999-774A-10

Alignment Scores:
Pred. No.: 1.63e-27 Length: 503
Score: 446.00 Matches: 140
Percent Similarity: 40.28% Conservative: 63
Best Local Similarity: 27.78% Mismatches: 193
Query Match: 16.52% Indels: 108
DB: 3 Gaps: 19

US-09-975-253-1 (1-1426) x US-08-999-774A-10 (1-503)

QY 62 CCACGGNTCTGT-----CCCTGGCTGTGTCTCGAGCTGGACCTGGGCAACTGGAGGC 115
Db 10 ProArgValLeuPheGlyGluTrpLeuLeuGlyGluIleSerSerGlyCysTyrGluGly 29
QY 116 GTGGCTGGGTGAACAAGAGCGCAGCGCTTCCGCAATCCCTTGGAGACAGCGCTACGG 175
Db 30 LeuGlnTrpLeuAspGluAlaArgThrCysPheArgValProTrpLysHisPheAlaArg 49
QY 176 CAGGATGCACAGCAGGAGGATTTCGGAATCTTCCAGGCTGGCCGAGGACCTGGTGCA 235
Db 50 LysAspLeuSerGluAlaAspAlaArgIlePheLysAlaTrpAlaValAlaArgGlyArg 69
QY 236 TATGTTTCC-----GGAGGGATAAGCCAGACCTGCCAAC----- 271
Db 70 TrpProProSerSerArgGlyGlyProProGluAlaGluThrAlaGluArgAla 89
QY 272 ---TGAAGAGGAATTCGCTCTGCTCCCAACCGCAAGAGGGTTCGTTTAGCAGAG 328
Db 90 GlyTrpLysThrAsnPheArgCysAlaLeuArgSerThrArgArgPheValMetLeuArg 109
QY 329 GACCGGACGACGAGCCCTCAGCAGCCACATAAAATCTACAGTTTGTG----- 376
Db 110 AspAsnSerGlyAspProAlaAspProHisLysValTyrAlaLeuSerArgGluLeuCys 129
QY 377 ---AACTCAGGAGTTGG-----GACTTTTCCAGCCAGACACCTCTCCGGAC 421
Db 130 TrpArgGluGlyProGlyThrAspGlnThrGluAlaGluAlaProAlaValPro 149
QY 422 ACCAATGGTGA----- 433
Db 150 ProGlnGlyGlyProProArgProPheLeuAlaHisThrProAlaGlyLeuGlnAlaPro 169
QY 434 -----GGCAGTACTTCTGAT-----ACCCAG 454
Db 170 GlyProLeuProAlaProAlaGlyAspLysGlyAspLeuLeuGlnAlaValGlnGln 189
QY 455 GAAGACATTCGTGATGAGTTACTG---GGTAACATGTGTGTGGCCCACTCCCA----- 505
Db 190 SerCysLeuAlaAspHisLeuLeuThrAlaSerTrpGlyAlaAspProValProThrLys 209
QY 506 -----GATCCGGGA 514
Db 210 AlaProGlyGluGlyGlnGluGlyLeuProLeuThrGlyAlaCysAlaGlyGlyProGly 229
QY 515 CCCCCA-----AGCTGGCTGTAGCCCTGAGCCCTGAGCCCTGCGCTCAG 553
Db 230 LeuProAlaGlyGluLeuTyrGlyTrpAlaValGluLysThrProSerProGlyProGln 249
QY 554 CCGCTGGGAGCCCCAGCTTGGACAATCCCACTCC-----TTCCCA 595
Db 250 ProAlaAlaLeuThrThrGlyGluAlaAlaProGluSerProHisGlnAlaGluPro 269
QY 596 AACCTGGGGCCCTCTGAGAACCCACTGAAGCGGTGTGTGGCGGGGGAAGAGTGGAG 655
Db 270 TyrLeuSerProSerProSerAlaCysThrAlaValGlnGluPro---SerProGlyAla 288
QY 656 TTCAGGTGACAGCCTTCTACCGGGCGCCCACTCTCCAGCAGACCATCTCTCTGCCCG 715
Db 289 LeuAspValThrIleMetTyrLysGlyArgThrValLeuGlnLysValValGlyHisPro 308
QY 716 GAGGCGCTGCGGTGGTGGG---TCCGAAGTGGGAGACAGCGCTGCTGGATGGCA 772
Db 309 SerCysThrPheLeuTyrGlyProProAspProAlaValArgAlaThrAspProGlnGln 328
QY 773 GTACACTGCCAGACCTTGGCATCTCCCTCAGACAGGGGAGTGATGAGTACGTAGG 832
Db 329 ValAlaPheProSerProAla---GluLeuProAspGlnLysGlnLeuArgTyrThrGlu 347
QY 833 CATGTCTGAGTGCCTGGGTGGGGGACTGCTCTCTGGCGGGCGGCGAGTGGCTCTGG 892
Db 348 GluLeuLeuArgHisValAlaProGlyLeuHisLeuGluLeuArgGlyProGlnLeuTrp 367

QY 893 GCCCAGCGCTGGGCACTGTCACACATATCTGGGCGAGTGCAGGAGCTCTCCCAAC 952
 DDb 368 AlaArgMetGlyLysCysLysValTyrTrpGluValGly-----381
 QY 953 AGCGGCGATGGCGCTGATGCGAG-----GTCCCAAG 985
 DDb 382 -----GlyProGlySerAlaSerProSerThrProAlaCysLeuLeuProArg 398
 QY 986 GACAGGAGGAGGCGTGTGACCTGGGCGCTTCAATGTAGATCTGATACCTTCACG 1045
 DDb 399 AsnCysAspThrProIlePheAspPheArgValPheGlnGluLeuValGluPheArg 418
 QY 1046 GAAGAAAGCGGAGCG---TCACACGCTATGCCCTCTGTTCTGTGGGAGTCATCG 1102
 DDb 419 AlaArgGlnArgGlySerProArgTyrThrIleTyrLeuGlyPheGlyGlnAspLeu 438
 QY 1103 CCCAGGACCGCGTGGACCAAGAGGCTGCTGATGCTCAAGTGTGTCGCGCTC 1162
 DDb 439 SerAlaGlyArgProLysGlyLysSerLeuValLeuValLysLeuGluProTrpLeuCys 458
 QY 1163 AGGGCGCTTGTAGAAATGCCCGGCTAGGGGTGCTCTCCCTGGAG---AATACTGTG 1219
 DDb 459 ArgValHisLeuGluGlyThrGlnArgGluGlyValSerSerLeuAspSerSerLeu 478
 QY 1220 GACCTGCATTTCCAAAGCCACCCACTCTCCCTCACCTCCGACGAGGCTAC 1279
 DDb 479 SerLeuCysLeuSerSerAlaSerLeu-----TyrAspAspIleGluCysPhe 495
 QY 1280 CTGCAGGACTTG 1291
 DDb 496 LeuMetGluLeu 499

RESULT 7

US-08-611-280-25
 ; Sequence 25, Application US/08611280
 ; Patent No. 5891666
 ; GENERAL INFORMATION:
 ; APPLICANT: Matsuyama, Toshifumi
 ; APPLICANT: Grossman, Alex
 ; APPLICANT: Richardson, Christopher D.
 ; TITLE OF INVENTION: NOVEL GENES ENCODING LSIRF POLYPEPTIDES
 ; NUMBER OF SEQUENCES: 25
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Angen Canada Inc.
 ; STREET: 6733 Mississauga Road, Suite 303
 ; CITY: Mississauga
 ; STATE: Ontario
 ; COUNTRY: Canada
 ; ZIP: L5N 6J8
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/611,280
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Oleski, Nancy A.
 ; REGISTRATION NUMBER: 34,688
 ; REFERENCE/DOCKET NUMBER: A-338A
 ; INFORMATION FOR SEQ ID NO: 25:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 450 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-611-280-25
 Alignment Scores: 1.32e-23 Length: 450
 Pred. No.:

Score: 397.00 Matches: 125
 Percent Similarity: 41.63% Conservative: 54
 Best Local Similarity: 29.07% Mismatches: 187
 Query Match: 14.71% Indels: 64
 Gaps: 16
 DB: 2
 US-09-975-253-1 (1-1426) x US-08-611-280-25 (1-450)
 QY 77 TGGTGTGTGTCGACGTGGAGCTGGGCAACTGGAGCGTGGCTGGTGTGACAAAGAGC 136
 DDb 27 TrpLeuIleAspGlnIleAspSerGlyLysTyrProGlyLeuValTrpGluAsnGluGlu 46
 QY 137 CGCAGCGCTTCCGATCCCTTGGAGCAGCGCCTACGCGAGAT---GCACAGCAGGAG 193
 DDb 47 LysSerIlePheArgIleProTrpLysHisAlaGlyLysGlnAspTyrAsnArgGluGlu 66
 QY 194 GATTTCGGAATCTTCAGCGCTGGCGGAGGCGCTGTGTGCATATGTTCCCGGAGGAGAT 253
 DDb 67 AspAlaAlaLeuPheLysAlaTrpAlaLeuPheLysGlyLysPheArgGluGlyLeuAsp 86
 QY 254 AAGCCAGACCTGCCCAACCTGGAGAGGAATTTCCGCTCTGCGCTCAACCGCAAAGAGG 313
 DDb 87 LysProAspProProThrTrpLysThrArgLeuArgCysAlaLeuAsnLysSerAsnAsp 106
 QY 314 TTGCGTTTAGCAGAGACCGGAGCAAG---GACCCCTCAGCAGCCACATATAATCTACGAG 370
 DDb 107 PheGluGluLeuValGluArgSerGlnLeuAspIleSerAspProTrpLysValTyrArg 126
 QY 371 TTTGTGAACCTAGGAGTTGGGAGCTTTCCAGCCAGACACACCTCTCCGACACCAATGGT 430
 DDb 127 IleValProGluGlyAlaLysLysGlyAlaLysGlnLeuThrLeuLysAspProGlnMet 146
 QY 431 GGAGGCGAGT-----ACTTCTGATACCCAGGAGACATTTCTGGATGAGTTACTGGGT 481
 DDb 147 SerMetSerHisProTyrThrMetThrThrProTyrProSerLeuProAlaGlnValHis 166
 QY 482 AACATGCTGTGGTGGCCCACTC-----CCAGATCCGGGAGCCC 517
 DDb 167 AsnTyrMetMetProProLeuAspArgSerTrpArgAspTyrValProAspGlnProHis 186
 QY 518 CCAAGCTGGCTGTAGCCCTGAGCCCTGAGCCCTGCCCTCAGCCCTG-----559
 DDb 187 ProGluIle-----ProTyrGlnCysProMetThrPheGlyProArgGlyHisHis 203
 QY 560 ---CGGAGCCCGAGCTTGGACAAT-----CCCACCTCCCTTCCCAACCTGGGG 604
 DDb 204 TrpGlnGlyProAlaCysGluAsnGlyCysGlnValThrGlyThrPheTyrAlaCysAla 223
 QY 605 CCCTCTGAGAAC-----CCACTGAAGCGGCTGTGTGGTGGCGGGAAGAG 649
 DDb 224 ProProGluSerGlnAlaProGlyValProThrGluProSerIleArgSerAlaGluAla 243
 QY 650 TGGAGTTC-----GAGGTGACAGCCTTCTACCGGGGCGCCAGGTCTTC 694
 DDb 244 LeuAlaPheSerAspCysArgLeuHisIleCysLeuTyrTyrArgGluIleLeuValLys 263
 QY 695 CAGCAGACCATCTCTCCCGGAGGCGCTCGCGTGTGGGTCCGAGTGGGAGACAGG 754
 DDb 264 GluLeuThrThrSerSerProGluGlyCysArgIleSerHisGlyHisThrTyrAspAla 283
 QY 755 ACCTGCTGATGGCCAGTCACACTGCCAGACCTGCGATGTCCTGCACACAGGCGGA 814
 DDb 284 SerAsnLeuAsp---GlnValLeuPheProTyrPro-----GluAspAsnGly 298
 QY 815 GTGATGAGCTACGTGAGGCGATGCTGCTGAGCTGCGCTGGGTGGGAGGACTGCTCTGGCGG 874
 DDb 299 GlnArgLysAsnIleGluLysLeuLeuSerHisLeuGluArgGlyValValLeuTrpMet 318
 QY 875 GCCGGGCGAGTGGCTCTGGGCGGCGGCGGCTGGGCGGACTGCCACACATCTATGGGCACTGAGC 934
 DDb 319 AlaProAspGlyLeuTyrAlaLysArgLeuCysGlnSerArgIleTyrTrp-----335
 QY 935 GAGGAGCTGCTCCCAACAGCGGCGATGGGCT-----GATGCG 973

Db 336 -----AspGlyProLeuAlaLeuCysAsnAspArgProAsn 347
 QY 974 GAGTCCCAAGACAGGAGGCGTGTTCAGCTGGGGCCCTTCATTGTAGATCTG 1033
 Db 348 LysLeuGluArgAspGlnThrCysLysLeuPheAspThrGlnGlnPheLeuSerGluLeu 367
 QY 1034 ATTACCTTCACGGAAGGACGACGCTACACGAGCTATGCCCTCTGGTTCCTGTGGGG 1093
 Db 368 GlnAlaPheAlaHisGlyArgSerLeuProArgPheGlnValThrLeuCysPheGly 387
 QY 1094 GAGTCATGGCCCCCAGGACGACGCTGGACCAAGAGGCTCGTATGGTCAAGGTTGTGCC 1153
 Db 388 GluGluPheProAspProGlnArg---GlnArgLysLeuIleThrAlaHisValGluPro 406
 QY 1154 ACCTGCTCAGGCGCTGTAGAAATGGCCCGGTAGGGGTGCCCTCCCTGGAGAAT 1213
 Db 407 LeuLeuAlaArgGlnLeuTyrrPheAlaGlnGlnAsnSerGlyHisPheLeuArgGly 426
 QY 1214 ACTGTGGACTG-----CACATTTCAC 1237
 Db 427 ---TyrAspLeuProGluHisIleSerAsn 435

RESULT 8

US-09-195-940-25
 ; Sequence 25, Application US/09195940
 ; Patent No. 6258935
 ; GENERAL INFORMATION:
 ; APPLICANT: Matsuyama, Toshifumi
 ; APPLICANT: Grossman, Alex
 ; APPLICANT: Richardson, Christopher D.
 ; TITLE OF INVENTION: NOVEL GENES ENCODING LSIRF POLYPEPTIDES
 ; NUMBER OF SEQUENCES: 25
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Amgen Canada Inc.
 ; STREET: 6733 Mississauga Road, Suite 303
 ; CITY: Mississauga
 ; STATE: Ontario
 ; COUNTRY: Canada
 ; ZIP: L5N 6J8
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; FILING DATE: US/09/195,940
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/611,280
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Oleski, Nancy A.
 ; REGISTRATION NUMBER: 34,688
 ; REFERENCE/DOCKET NUMBER: A-338A
 ; INFORMATION FOR SEQ ID NO: 25:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 450 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 US-09-195-940-25

Alignment Scores:
 Pred. No.: 1.32e-23 Length: 450
 Score: 397.00 Matches: 125
 Percent Similarity: 41.63% Conservative: 54
 Best Local Similarity: 29.07% Mismatches: 187
 Query Match: 14.71% Indels: 64
 DB: 3 Gaps: 16

US-09-975-253-1 (1-1426) x US-09-195-940-25 (1-450)
 QY 77 TGCTGTGTTCGAGCTGGAGCACTGGGGCGTGGCTGGGTGAACAAGAGC 136
 Db 27 TrpLeuIleAspGlnIleAspSerGlyLysTyrrProGlyLeuValTrpGluAsnGluGlu 46
 QY 137 CGACACGGCTTCGCTATCCCTTGGAAAGCAGCGGCTACGGCAGGAT---GCACAGCAGGAG 193
 Db 47 LysSerIlePheArgIleProTrpLysHisAlaGlyLysGlnAspTyrrAsnArgGluGlu 66
 QY 194 GATTTCGGATCTTCAGGCTGGGCGGAGGCACTGTCATATGTTCCCGGAGGGAT 253
 Db 67 AspAlaAlaLeuPheLysAlaTrpAlaLeuPheLysGlyLysPheArgGluGlyIleAsp 86
 QY 254 AAGCCAGACTGCCAACCTCGGAAGGAATTTCCGCTCTGCCCTCAACGCAAGAAAGGG 313
 Db 87 LysProAspProProThrTrpLysThrArgLeuArgCysAlaLeuAsnLysSerAsnAsp 106
 QY 314 TTGCGTTTACGAGAGGACCGGAGCAAG---GACCCCTACAGCCACCATAAATCTACGAG 370
 Db 107 PheGluGluLeuValGluArgSerGlnLeuAspIleSerAspProTyrrLysValTyrrArg 126
 QY 371 TTTGTGAACCTCAGGAGTTGGGACTTTTCCAGCCAGACACCTCTCCGACACCAATGTT 430
 Db 127 IleValProGluGlyAlaLysLysGlyAlaLysGlnLeuThrLeuGluAspProGlnMet 146
 QY 431 GGAGGCAGT-----ACTTCTGATACCCAGGAAGACATTTCTGGATCAGTTACTGGGT 481
 Db 147 SerMetSerHisProTyrrThrMetThrProTyrrProSerLeuProAlaGlnValHis 166
 QY 482 AACATGGTGTGGCCCTCCTC-----CCAGATCCGGGAGCC 517
 Db 167 AsnTyrrMetProLeuAspArgSerTrpArgAspTyrrValProAspGlnProHis 186
 QY 518 CCAAGCTGGGTAGCCCTCAGCCCTGAGCCCTCCCTCAGCCCTG-----559
 Db 187 ProGluIle-----ProTyrrGlnCysProMetThrPheGlyProArgGlyHisHis 203
 QY 560 ---CGAGCCCTCTGGACAAT-----CCACTCTCTCCCAACCTGGGG 604
 Db 204 TrpGlnGlyProAlaCysGluAsnGlyCysGlnValThrGlyThrPheTyrrAlaCysAla 223
 QY 605 CCCTCTGAGAAC-----CACTGAAGCGCTGTGGTCCGCGGGGGAAGAG 649
 Db 224 ProProGluSerGlnAlaProGlyValProThrGluProSerIleArgSerAlaGluAla 243
 QY 650 TGGGAGTTC-----GAGGTGACAGCTTCTACCGGGGCGCCCAAGTCTTC 694
 Db 244 LeuAlaPheSerAspCysArgLeuHisIleCysLeuTyrrArgGluIleLeuValLys 263
 QY 695 CAGCAGACCATCTCTGCTCCGAGGCGCTGGGCTGGTGGGTCCGAAAGTGGGACAGG 754
 Db 264 GluLeuThrThrSerSerProGluGlyCysArgIleSerHisGlyHisThrTyrrAspAla 283
 QY 755 ACGTGCCTGGATGCCAGTCACATGCCAGACCTGGCATGCTCCCTGACAGACAGGGA 814
 Db 284 SerAsnLeuAsp---GlnValLeuPheProTyrrPro-----GluAspAsnGly 298
 QY 815 GTGATGAGCTACGTGAGCGCATGCTGAGCTGCTGGGTGGGAGGAGTGGCTCTGGCGG 874
 Db 299 GlnArgLysAsnIleGluLysLeuLeuSerHisLeuGluArgGlyValValLeuTrpMet 318
 QY 875 GCGGGCAGTGGCTCTGGGCGCAGCGCTGGGCGACTGCCACACATACTGGCGCAGTGAGC 934
 Db 319 AlaProAspGlyLeuTyrrAlaLysArgLeuCysGlnSerArgIleTyrrTrp-----335
 QY 935 GAGGAGCTGCTCCCAACAGCGGCGCATGGGCT-----GATGGC 973
 Db 336 -----AspGlyProLeuAlaLeuCysAsnAspArgProAsn 347
 QY 974 GAGTCCCAAGACAGGAGGCGTGTTCAGCTGGGGCCCTTCATTGTAGATCTG 1033
 Db 348 LysLeuGluArgAspGlnThrCysLysLeuPheAspThrGlnGlnPheLeuSerGluLeu 367

US-09-975-253-1 (1-1426) x US-09-562-466-25 (1-450)

QY	77	TGCTGGTTCGACGTGCATCGCCGCCACTGGAGGCGTGCTGGTTCAACAAGC	136
		: : :	
		: : :	
DB	27	TipLeuIleAspGlnileAspSerGlyNystyrProGlyLeuValTrpGluasngluGlu	46

QY	1211	AATACTGTGACCTG-----CACATTTCAC	1237
Db	427	Gly---TyrAspLeuProGluHisIleSerAsn	436

RESULT 11

US-08-611-280-2

Sequence 2, Application US/08611280

Patent No. 5891666

GENERAL INFORMATION:

APPLICANT: Matsuyama, Toshifumi

APPLICANT: Grossman, Alex

APPLICANT: Richardson, Christopher D.

TITLE OF INVENTION: NOVEL GENES ENCODING LSIRF POLYPEPTIDES

NUMBER OF SEQUENCES: 25

CORRESPONDENCE ADDRESS:

ADDRESSEE: Amgen Canada Inc.

STREET: 6733 Mississauga Road, Suite 303

CITY: Mississauga

STATE: Ontario

COUNTRY: Canada

ZIP: L5N 6T8

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/611,280

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Oleski, Nancy A.

REGISTRATION NUMBER: 34,688

REFERENCE/DOCKET NUMBER: A-338A

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 450 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLSCULE TYPE: protein

US-08-611-280-2

Alignment Scores:	6.96e-23	Length:	450
Pred. No.:	388.00	Matches:	124
Score:	40.13%	Conservative:	55
Best Local Similarity:	27.80%	Mismatches:	183
Query Match:	14.38%	Indels:	84
DB:	2	Gaps:	17

US-09-975-253-1 (1-1426) x US-08-611-280-2 (1-450)

QY	77	TGCTGGTTCGACCTGGACCTGGGGGCGTGGGAGGCGTGGGTGAACAAGAGC	136
Db	27	TripleuileAspGlnIleAspSerGlyLysTyrProGlyLeuValTripleuAsnGlu	46
QY	137	CGACCGCGTTCGCGATCCCTTGGAGACGACGCGCTACGCCAGGAT---GCACAGCAGGAG	193
Db	47	LysSerValPheArgIleProTriPlyHisAlaGlyLysGlnAspTyrAsnArgGlu	66
QY	194	GATTTCGGAATCTCCAGGCGTGGCGGAGGCCACTGTGTGCATATGTTCCCGGAGGGAT	253
Db	67	AspAlaIleuPheLysAlaTriPalaLeuPheLysGlyLysPheArgGluGlyIleAsp	86
QY	254	AAGCCAGACCTGCCAATCTGGAGAGAAATTCGGCTCTGCCCTCAACCGCAAGAGGG	313
Db	87	LysProAspProProThrTriPlyThrArgLeuArgCysAlaLeuAsnLysSerAsnAsp	106
QY	314	TTCGGTTAGCAGAGGACCGGACGAAG---GACCCCTCACGACCCACATAAAATCTACGAG	370
Db	107	PheGluGluLeuValGluArgSerGlnIleSerAspProTriLysValTyrArg	126

; Sequence 2, Application US/09562466

Patent No. 6369202
 GENERAL INFORMATION:
 APPLICANT: Matsuyama, Toshifumi
 Grossman, Alex
 Richardson, Christopher D.
 TITLE OF INVENTION: NOVEL GENES ENCODING LSIRF POLYPEPTIDES
 NUMBER OF SEQUENCES: 25
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Amgen Canada Inc.
 STREET: 6733 Mississauga Road, Suite 303
 CITY: Mississauga
 STATE: Ontario
 COUNTRY: Canada
 ZIP: L5N 6J8
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/562,466
 FILING DATE: 01-May-2000
 CLASSIFICATION: <unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 09/195,940
 FILING DATE: <unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: Oleski, Nancy A.
 REGISTRATION NUMBER: 34,688
 REFERENCE/DOCKET NUMBER: A-338A
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 450 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 2:
 US-09-562-466-2

Alignment Scores:
 Pred. No.: 6.96e-23 Length: 450
 Score: 388.00 Matches: 124
 Percent Similarity: 40.13% Conservative: 55
 Best Local Similarity: 27.80% Mismatches: 183
 Query Match: 14.38% Indels: 84
 DB: 3 Gaps: 17

US-09-975-253-1 (1-1426) x US-09-562-466-2 (1-450)

QY	77	TGGCTGGTTCGAGCTGGAGCACTGGGCGCTGGCTGGTGACAAAGC	136
Db	27	TrpLeuLeuAspGlnIleAspSerGlyLysTyrProGlyLeuValTrpGluAsnGlu	46
QY	137	CGCAGCGCTTCGATCCCTTGAAGCAGCGCTACGGCAGGAT---GCACACGAGG	193
Db	47	LysSerValPheArgIleProTrpLysHisAlaGlyLysGlnAspTyrAsnArgGlu	66
QY	194	GATTTCGAATCTCCAGCTGGCGCCGAGCCACTGGTGCATATGTTCCCGGAGGAT	253
Db	67	AspAlaAlaLeuPheLysAlaTrpAlaLeuPheLysGlyLysPheArgGluGlyLeu	86
QY	254	AAGCCAGACCTGCCAACCTGGAGAGGAAATTCGCTCTGCCTCAACCCGCAAGAGG	313
Db	87	LysProAspProProThrTrpLysThrArgLeuArgCysAlaLeuAsnLysSerAsn	106
QY	314	TTGCTTTAGCAGACCGGAGCAG---GACCCTCAGCCACCATATAATCTACGAG	370
Db	107	PheGluGluLeuValGluArgSerGlnLeuAspIleSerAspProTyrLysValTyr	126
QY	371	TTGTG-----AACTCAGGAGTTGGGACTTTTCCAGCCAGACACCTCT---	415
Db	127	IleValProGluGlyAlaLysLysGlyAlaLysGlnLeuThrLeuAspAspThrGln	146

QY	416	-----COGACACCAATGGTGGAGCGAGTACTTCTGTATACCCAGAA	457
Db	147	AlaMetGlyHisProTyrProMetThrAlaProTyrGlySerLeuProAlaGlnGln	165
QY	458	GACATTCGTGGATGAGTTACTGGGTAAACATGGTGTGGCCCA	499
Db	166	-----ValHisAsnTyrMetMetProProHisAspArgSerTrpArg	179
QY	500	-----CTCCAGATCCGGGACCCCAAGCTGGTGTAGCCCTGTAGCCCTGCCTCAG	553
Db	180	AspTyrAlaProAspGlnSerHisProGluLeu-----ProTyrGlnCysProVal	196
QY	554	CCCCCTG-----CGGAGCCCCCAGCTGGACAAT	580
Db	197	ThrPheGlyProArgGlyHisHisTrpGlnGlyProSerCysGluAsnGlyCysGln	216
QY	581	CCCACCTCCCTTCCCAACCTGGGGCCCTCTGAGAAC-----CCACTGAAG	625
Db	217	ThrGlyThrPheTyrAlaCysAlaProProGluSerGlnAlaProGlyLysProIleGlu	236
QY	626	CGGCTGTTGGTCCGGGGGAA-----GAGTGGAGATTCGAGGTGACAGCC	670
Db	237	ProSerIleArgSerAlaGluAlaLeuAlaLeuSerAspCysArgLeuHisIleCysLeu	256
QY	671	TTCTACCGGGCGCGCAAGTCTTCCAGCAGACCATCTCTGCGGAGGCGCTGGCTG	730
Db	257	TyrTyrArgAspIleLeuValLysGluLeuThrThrSerProGluGlyCysArgIle	276
QY	731	GTGGGTCCGAAGTGGGACAGGAGCGTCTGGATGCCAGTGCACACTGCCAGACCT	790
Db	277	SerHisGlyHisThrTyrAspValSerAsnLeuAsp---GlnValLeuPheProTyrPro	295
QY	791	GGCATGCTCCTGACAGCAGGCGAGTATGATGAGTACGTGAGGATCTGTGAGTGCCTG	850
Db	296	-----AspAsnGlyGlnArgLysAsnIleGluLysLeuLeuSerHisLeu	311
QY	851	GGTGGGGAGCTGCTCTCTGGGGCGCGGAGTGTCTGGCCCGGAGGCGCTGGGCGAC	910
Db	312	GluArgGlyLeuValLeuTrpMetAlaProAspGlyLeuTyrAlaLysArgLeuCysGln	331
QY	911	TGCCACACATCTGGGCGAGTGGAGGAGCTCTCCCAACAGCGGCGATGGCCT---	967
Db	332	SerArgIleTyrTrp-----AspGlyProLeu	340
QY	968	-----GATGCGAGGTCCCAAGGACCAAGGAGGAGGCGGTGTGAC	1009
Db	341	AlaLeuCysSerAspArgProAsnLysLeuGluArgAspGlnThrCysLysLeuPheAsp	360
QY	1010	CTGGGGCCCTTCATTTAGATCTGATTACCTTACGGAAGGAGCGGACCTCACCCAGC	1069
Db	361	ThrGlnGlnPheLeuSerGluLeuGlnValPheAlaHisHisGlyArgProAlaProArg	380
QY	1070	TATGCCCTCTGGTCTGTGGGGAGTCTATGGCCAGGACCGCTGGACCAAGAGG	1129
Db	381	PheGlnValThrLeuCysPheGlyGluGluPheProAspProGlnArg---GlnArgLys	399
QY	1130	CTCCTGATGTCAGAGTTGTGCCACGCTGCTCAGGGCGCTGTGGTAGAAATGGCCGGTA	1189
Db	400	LeuIleThrAlaHisValGluProLeuLeuAlaArgGlnLeuTyrPheAlaGlnGln	419
QY	1190	GGGGGTGCTCTCCCTGGAG-----AATACTGTGAGCCTG	1225
Db	420	AsnThrGlyHisPheLeuArgGlyTyrGluLeuProGluHisValThrThrProAspTyr	439
QY	1226	CACATTCCACAGCCAC	1243
Db	440	HisArgSerLeuArgHis	445

RESULT 14
 US-09-230-371A-30
 ; Sequence 30, Application US/09230371A
 ; Patent No. 634856

```

; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohnzky, Roy A
; APPLICANT: Russo, James J
; APPLICANT: Edelman, Isidore S
; APPLICANT: Moore, Patrick S
; TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS SEQUENCES AND
; FILE REFERENCE: 45185-G-PCT-US
; CURRENT FILING DATE: 1999-11-17
; PRIOR FILING DATE: 1997-07-22
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Human
US-09-230-371A-30

Alignment Scores:
Pred. No.: 5,22e-20 Length: 425
Score: 352.00 Matches: 118
Percent Similarity: 42.44% Conservative: 56
Best Local Similarity: 28.78% Mismatches: 178
Query Match: 13.04% Indels: 58
DB: 3 Gaps: 14

US-09-975-253-1 (1-1426) x US-09-230-371A-30 (1-425)
QY 65 CGGNTCTCCCTGGTGTGTGGAGCTGGGACCTGGGCAACTGGAGGGCTGGCCCTGG 124
DB 9 ArgLeuArgGlnTrpLeuIleGluGlnIleAspSerSerMetTyrProGlyLeuIleTrp 28
QY 125 GTGAACAAGAGCGCGCGCTTCCGCATCCCTTGAAGCAGCGCTACGGCAGGATGCA 184
DB 29 GluAsnGluGlySerMetPheArgIleProTrpLysHisAlaGlyLysGlnAspTyr 48
QY 185 CAGCAGGAG--GATTTCGGAATCTTCCAGGCTGGCGGAGCCACTGGTCATATGTT 241
DB 49 AsnGlnGluValAlaSerIlePheLysAlaTrpAlaValPheLysGlyLysPheLys 68
QY 242 CCGGGAGGTAAGCCAGACCTGCCAAGCTGGAGAGGAATTCGGCTCTCCCTCAAC 301
DB 69 GluGly---AspLysAlaGluProAlaThrTrpLysThrArgLeuArgCysAlaLeuAsn 87
QY 302 CGAAAGAAGGTTGCGTTTAGCAGAGGACCGGAGCAAG--GACCCTCACGACCCACAT 358
DB 88 LysSerProAspPheGluGluValThrAspArgSerGlnLeuAspIleSerGluProTyr 107
QY 359 AAAATCTACGAGTTTGTG-----AATCAGGAGTTGGGACTTT 397
DB 108 LysValTyrArgIleValProGluGluAspGlnLysCysLysLeuGlyValAlaThrAla 127
QY 398 TCCAGCCAGACACCTCCGACACCAATGGTGAGGACGACTACTCTGAT-----ACC 451
DB 128 GlyCysValAsnGluValThrGluMetGluCysGlyArgSerGluIleAspGluLeuIle 147
QY 452 CAGGAAGACATCTGGATGACTTCTGGGTAAACATGTGGTGGCCCACTCCAGATCCG 511
DB 148 LysGluProSerValAspAspTyrMetGlyMetIleLysArgSerProSerPro----- 165
QY 512 GGACCCCAAGCTGGTGTAGCCCTCAGCCCTGAGCCCTCAGCCCTCGGAGGCCACG 571
DB 166 -----ProAspAlaCysArgSerGlnLeu----- 173
QY 572 TTGGACAATCCCACTCCCTCCCAACCTGGGGCCCTCTGAGAACCCACTGAAGCGGCTG 631
DB 174 -----LeuProAspTrpTrpAlaHisGluProSerThrGlyArg 187
QY 632 TTGGTCCGGGGGAAGAGTGG-----GAGTTCGAGGTG 664
DB -----

```

```

DB 188 LeuValThrGlyTyrThrThrTyrAspAlaHisHisSerAlaPheSerGlnMetValIle 207
QY 665 ACAGCCTTCTACCGGGCGCCAAAGTCTTCCACAGCAGCATCTCTGCGCGGAGGCGCTG 724
DB 208 SerPheTyrTyrGlyLysLeuValGlyGlnAlaThrThrThrCysProGluGlyCys 227
QY 725 CGGCTGGTGGGTCCGAAGTGGGAGACAGGACGCTCCCTGGATGGCCAGTCACACTGCCA 784
DB 228 ArgLeuSerLeuSerGlnProGly-----LeuProGlyThrLysLeuTyrGlyPro 244
QY 785 GACCTGGCATGTCCTG-----ACACAGAGGGAGTGTAGCTACGTG 829
DB 245 Glu---GlyLeuGluLeuValArgPheProProAlaAspThrIleProSerGluArgGln 263
QY 830 AGGCATGTCTGAGCTGCTGGTGGGACTG-----GCTCTGCGGGCGCGG 880
DB 264 ArgGlnValThrArgLysLeuPheGlyHisLeuGluArgGlyValLeuHisSerSerArg 283
QY 881 CAGTGGCTCTGGGCCCGAGCGGTGGGGCACTGCCACATACTGGGCACGTGAGCGAGGAG 940
DB 284 GlnGlyValPheValLysArgLeuCysGlnGlyArgValPheCysValValValVal 303
QY 941 CTGCTCCCCAACAGCGGCGATGGGCTGATGGCGAGGTCCCCAAGACAAAGAGGAGGC 1000
DB 304 ValCys-----LysGlyArgProAsnLysLeuGluArgAspGluValValGln 319
QY 1001 GTGTTTGACCTGGGCGCCTTCACTGTAGATCTGATTACCTTCACGGAAGAGCGGACGC 1060
DB 320 ValPheAspThrSerGlnPheArgGluLeuGlnGlnPheTyrAsnSerGlnGlyArg 339
QY 1061 TCACACAGCTATGCTCTGTGTGTGTGGGGAGTCAATGGCCCCCAGGACCGCGTGG 1120
DB 340 LeuProAspGlyArgValValLeuLysCysPheGlyGluGluPheProAspMetAlaProLeu 359
QY 1121 ACCAAGAGCTCGTATGTCAGGTTCAGGTTGTGCCACGTCGCTCAGGCGCTTGTAGAATG 1180
DB 360 ArgSerLysLeuIleLeuValGlnIleGluGlnLeuTyrValArgGlnLeuAlaGlu 379
QY 1181 GCC---CGGTAGGGGGTGCCTCTCTCTCTG 1207
DB 380 AlaGlyLysSerCysGlyAlaGlySerVal 389

RESULT 15
US-09-230-371A-29
; Sequence 29, Application US/09230371A
; Patent No. 6348586
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohnzky, Roy A
; APPLICANT: Russo, James J
; APPLICANT: Edelman, Isidore S
; APPLICANT: Moore, Patrick S
; TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS SEQUENCES AND
; FILE REFERENCE: 45185-G-PCT-US
; CURRENT FILING DATE: 1999-11-17
; PRIOR FILING DATE: 1997-07-22
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29
; LENGTH: 393
; TYPE: PRT
; ORGANISM: Human
US-09-230-371A-29

Alignment Scores:
Pred. No.: 5,65e-16 Length: 393
Score: 301.50 Matches: 108
Percent Similarity: 41.46% Conservative: 62
Best Local Similarity: 26.34% Mismatches: 177
Query Match: 11.17% Indels: 63

```

DB: 3 Gaps: 16
US-09-975-253-1 (1-1426) x US-09-230-371A-29 (1-393)
Qy 38 GCGCGACCATGGGAACCCCAAGCAGCGTCTCCCTGGTGGTGCAGCTGGAC 97
Dy 4 GlyArgAlaArgCysThrArgLysLeuArgAsn-----TTPValValGluValGlu 21
Qy 98 CTGGGCAACTGGAGGCGTGGCTGGTGAACAGCCGACGCGCTTCCGATCCCT 157
Dy 22 SerGlyGlnPheProGlyValCysTrpAspThrAlaLysThrMetPheArgIlePro 41
Qy 158 TGAAGCAGCGCTACCGCAGAT---GCACAGCAGGAGGATTTCCGAACTTCCAGGCC 214
Dy 42 TrpLysHisAlaGlyLysGlnAspPheArgGluAspGlnAspAlaAlaPhePheLysAla 61
Qy 215 TGGCGCGAGGCCACTGGTGCAATCTTCCCGGAGGGATAAGCCAGACCTGCCAACCTGG 274
Dy 62 TrpAlaIlePheLysGlyLysTyLysGluGly---AspThrGlyGlyProAlaValTrp 80
Qy 275 AAGAGGAATTTCCGCTCTGCCCTCAACCGCAAGAGGGTTCCGTTTAGCAGAGGACCGG 334
Dy 81 LysThrArgLeuArgCysAlaLeuAsnLysSerSerGluPheLysGluValProGluArg 100
Qy 335 AGCAAG---GACCCTCAGACCCACCAATAAATCTACGAGTTTGTGNACTCAGAGTTGGG 391
Dy 101 GlyArgMetAspValAlaGluProTyLysValTyLysGlnLeuLeuProProGlyIleVal 120
Qy 392 GACTTTTCCAGCCAGACCTCT-----CCGACACCAATGGTGGAGGAGTACTTCT 445
Dy 121 Ser---GlyGlnProGlyThrGlnLysValProSerLysArgGlnHisSerSerValSer 139
Qy 446 GATACCCAGCAGACATCTGGATGAGTTACTGGGTAAACATGTGTGGCCCCACTC--- 502
Dy 140 SerGluArgLysGlu---GluGluAspAlaMetGlnAsnCysThrLeuSerProSerVal 158
Qy 503 -----CCAGAT 508
Dy 159 LeuGlnAspSerLeuAsnAsnGluGluGluGlyAlaSerGlyGlyAlaValHisSerAsp 178
Qy 509 CCGGACCCCAAGCTGGCTGTAGCCCTGAGCCCTGCCCTCAGCCCTGCCGAGCCGCC 568
Dy 179 IleGlySerSerSerSerSerProGluPro-----Gln 191
Qy 569 AGCTTGGACAACTCCACTCCCTCCCAACCTGGGGCCCTCTGAGAACCCACTGAAGCGG 628
Dy 192 GluValThrAspThrThrGluAlaProPheGlnGly---AspGlnArgSerLeuGluPhe 210
Qy 629 CTGTTGGTGGCGGGAAGAGTGGAGTTCCAGGTGACAGCTTCTACCGGGCGCCGCAA 688
Dy 211 LeuLeuProGluProAspTyLysSerLeuLeuLeuThrPheIleTyLysAsnGlyArgVal 230
Qy 689 GTCTTCCAG-----CAGACCATCTCTGCCCGGAGGCGCTGCGGCTGGTGGGTCC 739
Dy 231 ValGlyGluAlaGlnValGlnSerLeuAspCys-----ArgLeuValAlaGlu 246
Qy 740 GAAGTGGGAGCAGCAGCTGCTGGATGGCCAGTCACTGCGCAGACCTGGCATGTCC 799
Dy 247 ProSerGlySerGluSerMetGlu---GlnValLeuPheProLysPro----- 262
Qy 800 CTGACACAGGAGGAGTATGATGAGTACGTGAGGAGTGTGCTGAGCTGCTGGTGGGGA 859
Dy 263 -----GlyProLeuGluProThrGlnArgLeuLeuSerGlnLeuGluArgGly 278
Qy 860 CTGGCTCTCTGGCGCGCGGCGAGTGGCTGTGGGCCAGCGCTGGGCACTGCCACACA 919
Dy 279 IleLeuValAlaSerAsnProArgGlyLeuPheValGlnArgLeuCysProIleProIle 298
Qy 920 TACTGGCAGTGGAGGAGCTGTCTCCCAACCGCGGATGGCCCTGATGGCGAG--- 976
Dy 299 SerTrp-----AsnAlaProGlnAlaProProGlyProGly 310
Qy 977 -----GTCCCAAGCAAGGAGGAGCGTGTGTGACCTGGGCGCTTCATTGTA 1027

Db 311 ProHisLeuLeuProSerAsnGluCysValGluLeuPheArgThrAlaTyrPheCysArg 330
Qy 1028 GATCTGATTACCTTCACGGAAGGAGCGCTCACACGCTATGCCCTCTGTTCTGT 1087
Dy 331 AspLeuValArgTyrPheGlnGlyLeuGlyProProLysPheGlnValThrLeuAsn 350
Qy 1088 GTGGGGGAGTTCATGGCCCCCAGGACAGCCGCTGGACCAAGAGCTCGTATCGTCAAGTT 1147
Dy 351 PheTrpGluGluSerHisGlySerHisThrProGlnAsnLeuIleThrValLysMet 370
Qy 1148 GTGCCACAGTCCCTCAGGCGCTTGGTAGAA 1177
Dy 371 GluGlnAlaPheAlaArgTyrLeuLeuGlu 380
Search completed: October 28, 2004, 15:44:18
Job time : 57 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 28, 2004, 15:15:00 ; Search time 149 Seconds
(without alignments)
6205.795 Million cell updates/sec

Title: US-09-975-253-1

Perfect score: 2699

Sequence: 1 ggttcagctgcgcgaagc.....aaaaaaaaaaaaaaaaaaaa 1426

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1370721 seqs, 324215800 residues

Total number of hits satisfying chosen parameters: 2741442

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlpx
-Q=/cgn2_1/USPTO spoal_p/US09975253/runat_28102004_111511_19667/app_query.fasta_1.1607
-DB=Published Applications AA -QPMF=fastan -SUFFIX=n2p.rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITFS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40 cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09975253 @CGN_1_1_224 @runat_28102004_111511_19667
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBIOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-XGAPOP=6 -XGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELETE=7

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result Query

No.	Score	Match	Length	DB	ID	Description
1	2326	86.2	427	9	US-09-975-253-2	Sequence 2, Appli
2	2326	86.2	427	14	US-10-453-478-16	Sequence 16, Appli
3	2018	74.8	392	15	US-10-276-774-2333	Sequence 31, Appli
4	399.5	14.8	94	15	US-10-258-666-31	Sequence 2929, Ap
5	257	9.5	395	15	US-10-264-049-2929	Sequence 3, Appli
c	247.5	9.2	19695	15	US-10-084-846A-3	Sequence 374, App
6	246	9.1	373	10	US-09-809-391-374	Sequence 374, App
7	246	9.1	373	10	US-09-882-171-374	Sequence 374, App
8	246	9.1	373	14	US-10-164-861-374	Sequence 7, Appli
9	246	9.1	373	14	US-10-084-846A-7	Sequence 46, Appli
c	243.5	8.9	19652	15	US-09-805-020-46	Sequence 8, Appli
11	240	8.9	292	9	US-10-084-846A-6	Sequence 356, App
c	239.5	8.9	19662	15	US-10-084-846A-8	Sequence 1897, Ap
c	234.5	8.8	19608	15	US-10-755-889-356	Sequence 5, Appli
c	231.5	8.6	164	16	US-10-276-774-1897	Sequence 141510,
c	222.5	8.3	19723	15	US-10-084-846A-5	Sequence 40, Appli
c	222	8.3	19723	15	US-10-084-846A-5	Sequence 40, Appli
c	219.5	8.2	405	16	US-10-437-963-141510	Sequence 30, Appli
c	219.5	8.1	733	15	US-10-138-588-40	Sequence 30, Appli
c	218	8.1	720	14	US-10-342-331-4	Sequence 30, Appli
c	214.5	7.9	744	14	US-10-171-311-40	Sequence 30, Appli
c	214.5	7.9	744	14	US-10-236-055A-30	Sequence 30, Appli
c	214.5	7.9	744	14	US-10-301-822-37	Sequence 30, Appli
c	213	8.0	960	14	US-10-342-331-5	Sequence 30, Appli
c	213	8.0	19725	15	US-10-084-846A-4	Sequence 4, Appli
c	211.5	7.9	408	16	US-10-437-963-184023	Sequence 184023,
c	211.5	7.9	504	14	US-10-342-331-3	Sequence 3, Appli
c	211.5	7.8	744	14	US-10-301-822-39	Sequence 39, Appli
c	210.5	7.9	466	16	US-10-437-963-167964	Sequence 167964,
c	209	7.8	1366	15	US-10-402-089-10	Sequence 10, Appli
c	209	7.8	1366	15	US-10-402-072A-10	Sequence 10, Appli
c	208.5	7.8	336	15	US-10-425-114-56601	Sequence 56601, A
c	208.5	7.8	903	14	US-10-156-761-11093	Sequence 11093, A
c	208.5	7.8	1078	14	US-10-058-124-21	Sequence 21, Appli
c	208	7.8	1014	14	US-10-194-441A-1	Sequence 1, Appli
c	207.5	7.8	780	15	US-10-441-965-24	Sequence 24, Appli
c	207.5	7.8	837	15	US-10-096-986-103	Sequence 103, App
c	207	7.7	555	16	US-10-437-963-203630	Sequence 203630,
c	207	7.7	1057	15	US-10-104-889-16	Sequence 16, Appli
c	207	7.7	1057	15	US-10-104-889-20	Sequence 20, Appli
c	207	7.7	1107	15	US-10-104-889-11	Sequence 11, Appli
c	207	7.7	1169	15	US-10-104-889-6	Sequence 6, Appli
c	207	7.7	1171	15	US-10-104-889-8	Sequence 8, Appli
c	207	7.7	1388	15	US-10-104-889-10	Sequence 10, Appli
c	207	7.7	1461	16	US-10-468-091-25	Sequence 25, Appli
c	207	7.7	1464	10	US-09-918-715-261	Sequence 261, App

ALIGNMENTS

RESULT 1

US-09-975-253-2
; Sequence 2, Application US/09975253
; Patent No. US2002016494A1
; GENERAL INFORMATION:
; APPLICANT: Paul Moore et al.
; TITLE OF INVENTION: Methods and Compositions for Treating and Preventing Infection
; FILE REFERENCE: PFI96P1
; CURRENT FILING DATE: 2001-10-12
; PRIOR FILING DATE: 60/239,963
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 427
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (8)

OTHER INFORMATION: Xaa equals Ile, Leu, Phe, or Val
US-09-975-253-2

Alignment Scores:
Pred. No.: 6, 47e-145 Length: 427
Score: 2326.00 Matches: 427
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 86.18% Indels: 0
DB: 9 Gaps: 0

US-09-975-253-1 (1-1426) x US-09-975-253-2 (1-427)

QY	47	ATGGGAACCCCAAAACCCAGGNTCTGCTGGCTGGTGTGCTGACCTGGGGCAA	106
DB	1	MetGlyThrProLysProArg***LeuProTrpLeuValSerGlnLeuAspLeuGln	20
QY	107	CTGGAGGGCTGGCTGGTGAACAGAGCCGACGCTTCCGCATCCCTTGGAGCAC	166
DB	21	LeuGluGlyValAlaTrpValAsnLysSerArgTrpArgPheArgLleProTrpLysHis	40
QY	167	GGCTACGCGCAGATGACAGCAGGAGGATTTCCGAATCTTCCAGGCTGGGCCGAGGCC	226
DB	41	GlyLeuArgGlnAspAlaGlnGlnGluAspPheGlyIlePheGlnAlaTrpAlaGluAla	60
QY	227	ACTGTGTCATATGTTCCCGGGAGGGATAGCCAGACCTGCCAACCTGGAGAGGATTTTC	286
DB	61	ThrGlyAlaTrpValProGluArgAspLysProAspLeuProThrTrpLysArgAsnPhe	80
QY	287	CGCTGCTGCCCTCAACCGCAAGAGGTTGCGTTTAGCAGAGGACCGGAGCAAGACCTG	346
DB	81	ArgSerAlaLeuAsnArgLysGluGlyLeuArgLeuAlaGlnAspArgSerLysAspPro	100
QY	347	CACGACCCACATAAAATCTACAGATTGTGAATCAGGAGTTGGGACTTTCCAGCCA	406
DB	101	HisAspProHisLysIleTrpGluPheValAsnSerGlyValGlyAspPheSerGlnPro	120
QY	407	GACACTCTCCGACACCAATGTTGGAGGAGTACTTCTGATACCCAGGAGACATCTG	466
DB	121	AspThrSerProAspThrAsnGlyGlySerThrSerAspThrGlnGluAspIleLeu	140
QY	467	GATGACTTACTGGTAACATGTTGGCCCACTCCAGATCCCGGACCCCAAGCCTG	526
DB	141	AspGluLeuGlyAsnMetValLeuAlaProLeuProAspProGlyProProSerLeu	160
QY	527	GCTGTAGCCCTCAGCCCTGCTCAGCCCTCGGAGCCCTTGGACATCCCACT	586
DB	161	AlaValAlaProGluProCysProGlnProLeuArgSerProSerLeuAspAsnProThr	180
QY	587	CCCTTCCCAACCTGGGCGCTCTGAGAACCCACTGAAGCGCTGTTGGTCGGGGGA	646
DB	181	ProPheProAsnLeuGlyProSerGluAsnProLeuLysArgLeuLeuValProGlyGlu	200
QY	647	GAGTGGGAGTTGAGGTGACAGCTTCTACGGGGCGCCCAAGTCTTCCAGCAGACCATC	706
DB	201	GluTrpGluPheGluValThrAlaPheTyArgGlyArgGlnValPheGlnGlnThrIle	220
QY	707	TCCTGCTCCGAGGCGCTGCGGCTGGTGGGTCCGAACTGGGAGACAGGACGCTGCTGGA	766
DB	221	SerCysProGluGlyLeuArgLeuValGlySerGluValGlyAspArgThrLeuProGly	240
QY	767	TGGCCAGTCACTGCCAGACCTTGGGATGTCCTGACACAGAGGAGTATGATGCTAC	826
DB	241	TrpProValThrLeuProAspProGlyMetSerLeuThrAspArgGlyValMetSerTrp	260
QY	827	GTGAGGATGTGCTGAGCTGCTGGGTGGGAGTCTCTCTGCGGGCGGGCAGTGG	886
DB	261	ValArgHisValLeuSerCysLeuGlyGlyGlyLeuAlaLeuTrpAlaGlyGlnTrp	280
QY	887	CTCTGGGCGCCAGCGCTGGGCACTGCCACATCTGCGGAGTGGAGGAGGAGCTGCTC	946
DB	281	LeuTrpAlaGlnArgLeuGlyHisCysHisThrTrpAlaValSerGluGluLeuLeu	300

QY	947	CCCAACAGCGGGCATGGGCTGATGGGAGGTCCCCAAGGACAAGAGGAGGCGTGT	1006
DB	301	ProAsnSerGlyHisGlyProAspGlyGluValProLysAspLysGluGlyValPhe	320
QY	1007	GACCTGGGCGCTTCAATTGATCTGATACCTTACGGAAGGAGGAGCGGCTCACCA	1066
DB	321	AspLeuGlyProPheIleValAspLeuIleThrPheThrGluGlySerGlyArgSerPro	340
QY	1067	CGCTATGCGCTCTGTTCTGTGGGGAGTCAAGGCCCGCCAGGACCGGCTGGACCAAG	1126
DB	341	ArgTyAlaLeuTrpPheCysValGlyGluSerTrpProGlnAspGlnProTrpThrLys	360
QY	1127	AGGCTCGTGTGTCAGGTTGTGCCCCACGTGCTCCAGGCGCTTGGTAGAATGCCCCGG	1186
DB	361	ArgLeuValMetValLysValValProThrCysLeuArgAlaLeuValGluMetAlaArg	380
QY	1187	GTAGGGGTGCTCTCTCCCTGGAGATACTGTGGACCTGCGCATTTCCAAAGCCACCCA	1246
DB	381	ValGlyGlyAlaSerSerLeuGluAsnThrValAspLeuHisIleSerAsnSerHisPro	400
QY	1247	CTCTCCCTCAGCTCCGACGAGTACAAAGCGCTTACCTGCGAGGACTTGGTGGAGGCGAT	1306
DB	401	LeuSerLeuThrSerAspGlnTyLysAlaTyLeuGlnAspLeuValGluGlyMetAsp	420
QY	1307	TTCCAGGCGCTGGGGAGAGC 1327	
DB	421	PheGlnGlyProGlyGluSer 427	
RESULT 2			
US-10-453-478-16			
; Sequence 16, Application US/10453478			
; Publication No. US20030208043A1			
; GENERAL INFORMATION:			
; APPLICANT: Paul Moore, Reiner Gentz, Hongjin Ji,			
; TITLE OF INVENTION: Human Genes, Sequences and			
; Expression Products			
; NUMBER OF SEQUENCES: 22			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: CABELLA, BYRNE, BAIN, GILFILLAN,			
; STREET: 6 BECKER FARM ROAD			
; CITY: ROSELAND			
; STATE: NEW JERSEY			
; COUNTRY: USA			
; ZIP: 07068			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: 3.5 INCH DISKETTE			
; COMPUTER: IBM PS/2			
; OPERATING SYSTEM: MS-DOS			
; SOFTWARE: WORD PERFECT 5.1			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/10/453,478			
; FILING DATE: 04-Jun-2003			
; CLASSIFICATION: 536			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: US/08/705,771			
; FILING DATE: August 30, 1996			
; ATTORNEY/AGENT INFORMATION:			
; NAME: MULLINS, J.G.			
; REGISTRATION NUMBER: 33,073			
; REFERENCE/DOCKET NUMBER: 325800-346 (PF196)			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: 973-994-1700			
; TELEFAX: 973-994-1744			
; INFORMATION FOR SEQ ID NO: 16:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 427 amino acids			
; TYPE: amino acid			
; TOPOLOGY: linear			
; MOLECULE TYPE: protein			
; SEQUENCE DESCRIPTION: SEQ ID NO: 16:			
US-10-453-478-16			

Alignment Scores:
Pred. No.: 6,47e-145 Length: 427
Score: 2326.00 Matches: 427
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 96.18% Indels: 0
DB: 14 Gaps: 0

US-09-975-253-1 (1-1426) x US-10-453-478-16 (1-427)

QY 47 ATGGGAACCCAAAGCCAGGNTCTGCTCGCTGGTGTGTCAGCTGGACCTGGGGCAA 106
DB 1 MetGlyThrProLysProArg***LeuProTrpLeuValSerGlnLeuAspLeuGln 20
QY 107 CTGAGGCGCTGGCTGGTGAACAGAGCCGACGCTCCCGATCCCTTTGGAAGCAC 166
DB 21 LeuGluGlyValAlaTrpValAsnLysSerArgThrArgPheArgIleProTrpLysHis 40
QY 167 GGCCTACGGCAGGATGCACAGCAGGAGGATTTCGGAATCTTCCAGGCTGGGCCGAGGCC 226
DB 41 GlyLeuArgGlnAspAlaGlnGlnGluAspPheGlyIlePheGlnAlaTrpAlaGluAla 60
QY 227 ACTGTGCATATGTTCCCGGAGGATAGCCAGACCTGCCAACCTGGAAGAGGAATTC 286
DB 61 ThrGlyAlaTyrValProGlyArgAspLysProAspLeuProThrTrpLysArgAsnPhe 80
QY 287 CGTCTGCCCTCAACCCGAAAGAGGTTGGCTTTAGCAGAGGACCGGAGCAAGACCCCT 346
DB 81 ArgSerAlaLeuAsnArgLysGluGlyLeuArgLeuAlaGluAspArgSerLysAspPro 100
QY 347 CACGACCCACATAAATCTACGAGTTGTGTAAGTTCAGGAGTTGGGACTTTTCCAGGCA 406
DB 101 HisAspProHisLysIleTyrGluPheValAsnSerGlyValGlyAspPheSerGlnPro 120
QY 407 GACACCTCTCCGGACCAATAGTGTGAGCGAGTACTTCTGATACCCAGGAGACATCTCTG 466
DB 121 AspThrSerProAspThrAsnGlyGlySerThrSerAspThrGlnGluAspLeu 140
QY 467 GATGAGTTACTGGGTAACATGTTGTGCCCCACTCCAGATCCGGGACCCCCAACGCTG 526
DB 141 AspGluLeuLeuGlyAsnMetValLeuAlaProLeuProAspProGlyProProSerLeu 160
QY 527 GCTGTAGCCCTGAGCCCTGCTCAGCCCTGCGGAGCCCGAGCTTGGACATCCCACT 586
DB 161 AlaValAlaProGluProCysProGlnProLeuArgSerProSerLeuAspAsnProThr 180
QY 587 CCCTTCCCAACCTGGGGCCCTCTGAGAACCCACTGGAAGCGGCTGTGGTGGCGGGGAA 646
DB 181 ProPheProAsnLeuGlyProSerGluAsnProLeuLysArgLeuLeuValProGlyGlu 200
QY 647 GAGTGGGAGTTGAGGTGACAGCCTTCTACCGGGGCCGCCAAGTCTTCCAGCAGACCATC 706
DB 201 GluTrpGluPheGluValThrAlaPheTyrArgGlyArgGlnValPheGlnGlnThrIle 220
QY 707 TCCTGCCCGGAGGCGCTCGGCTGTGGGTCCGAGTGGGAGCAGCAGCAGCTGCTCGA 766
DB 221 SerCysProGluGlyLeuArgLeuValGlySerGluValGlyAspArgThrLeuProGly 240
QY 767 TGGCCAGTCACACTGCCAGACCTGGCATGTCCCTGACAGACAGGGGAGTGATGAGTAC 826
DB 241 TrpProValThrLeuProAspProGlyMetSerLeuThrAspArgGlyValMetSerTyr 260
QY 827 GTGAGGCAATGTCTGAGTGTGGTGGGGAGCTGGCTCTCTGCGGGCCCGGCGAGTGG 886
DB 261 ValArgHisValLeuSerCysLeuGlyGlyLeuAlaLeuTrpArgAlaGlyGlnTrp 280
QY 887 CTCTGGGCCCCAGCGCTGGGCACTGCCACATCTGGCAGTGTGAGCAGGAGCTGCTC 946
DB 281 LeuTrpAlaGlnArgLeuGlyHisCysHisThrTyrTrpAlaValSerGluGluLeu 300
QY 947 CCCAACAGGGGCATGGGCTGTGATGGCGAGGTCCCAAGGACCAAGGAAGGAGGCTGTTT 1006

DB 301 ProAsnSerGlyHisGlyProAspGlyGluValProLysAspLysGluGlyValPhe 320
QY 1007 GACCTGGGGCCCTTCATTGTAGATCTGATTACCTTCACGGAAGGAGCGGACGCTCACCA 1066
DB 321 AspLeuGlyProPheIleValAspLeuIleThrPheThrGluGlySerGlyArgSerPro 340
QY 1067 CGCTATGCCCTCTGTTCTGTGGGGGAGTCATGGCCCCCAGGACCCGCTGGACCAAG 1126
DB 341 ArgTyrAlaLeuTrpPheCysValGlyGluSerTrpProGlnAspGlnProTrpThrLys 360
QY 1127 AGGCTCGTGATGGTCAAGTTGTGCCACGCGCTCAGGGCTGTGTAAGAAATGCCCGG 1186
DB 361 ArgLeuValMetValLysValValProThrCysLeuArgAlaLeuValGluMetAlaArg 380
QY 1187 GTAGGGGTGGCTCTCTCCCTGGAGAAATCTGTGGACCTGCACATTTCCAACAGCCACCA 1246
DB 381 ValGlyGlyAlaSerSerSerLeuGluAsnThrValAspLeuHisIleSerAsnSerHisPro 400
QY 1247 CTCTCCCTCACCTCGGACAGTACAGGCCTACCTGCGAGGACTTGGTGGAGGCGATGGAT 1306
DB 401 LeuSerLeuThrSerAspGlnTyrLysAlaTyrLeuGlnAspLeuValGluGlyMetAsp 420
QY 1307 TTCCAGGGCCCTGGGGAGAGC 1327
DB 421 PheGlnGlyProGlyGluSer 427

RESULT 3
US-10-276-774-2333
; Sequence 2333, Application US/10276774
; Publication No. US20040053245A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: No. US20040053245A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-030
; CURRENT APPLICATION NUMBER: US/10/276,774
; PRIOR FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 2700
; SOFTWARE: Custom
; SEQ ID NO 2333
; LENGTH: 392
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-276-774-2333

Alignment Scores:
Pred. No.: 1.25e-124 Length: 392
Score: 2018.00 Matches: 372
Percent Similarity: 98.41% Conservative: 0
Best Local Similarity: 98.41% Mismatches: 0
Query Match: 74.77% Indels: 6
DB: 15 Gaps: 1

US-09-975-253-1 (1-1426) x US-10-276-774-2333 (1-392)

QY 212 GCCTGGGCGGAGGACCTGCTGGTGCATATGTTCCCGGAGGGATAGCCAGACCTGCCAACC 271
DB 15 AlaTrpAlaGluAlaThrGlyAlaTyrValProGlyArgAspLysProAspLeuProThr 34
QY 272 TGGAGAGGAAATTTCCGCTCTGCCCTCAACCGCAAGAGGGTTGCGTTTAGCAGAGGAC 331
DB 35 TrpLysArgAsnPheArgSerAlaLeuAsnArgLysGluGlyLeuArgLeuAlaGluAsp 54
QY 332 CGGAGCAAGGACCTCTACGACCCACATAAAATCTTACGAGTTTGTGAATCTCAGGAGTTGG 391
DB 55 ArgSerLysAspProHisAspProHisLysIleTyrGluPheValAsnSerGlyValGly 74
QY 392 GACTTTTCCGACGACAGACCTCTCCCGACACCAATGGTGGAGGAGCTACTCTTGATACC 451

FILE REFERENCE: 082382-0000000US
CURRENT APPLICATION NUMBER: US/10/258,666
CURRENT FILING DATE: 2003-05-23
PRIOR APPLICATION NUMBER: JP 2000-126741
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: WO PCT/JP01/03700
PRIOR FILING DATE: 2001-04-27
NUMBER OF SEQ ID NOS: 42
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 31
LENGTH: 94
TYPE: PRT
ORGANISM: Rattus norvegicus
FEATURE:
OTHER INFORMATION: RHDH-279-1 ORF
US-10-258-666-31

Alignment Scores:
Pred. No.: 4,42e-18 Length: 94
Score: 399.50 Matches: 77
Percent Similarity: 87.23% Conservative: 5
Best Local Similarity: 81.91% Mismatches: 11
Query Match: 14.80% Indels: 1
DB: 15 Gaps: 1

US-09-975-253-1 (1-1426) x US-10-258-666-31 (1-94)

QY 1046 GAAGAAAGGAGCGCTACACGCTATGCCCTCTGTTCTGTTGGGGAGTCATGGCCC 1105
Db 2 GluGlySerArgHisSerProArgTyrThrLeuPheCysValGlyGluSerTrpPro 21
QY 1106 CAGGACCGCGCTGGACCAAGAGGCTCGTGTAGTGTCAAGTGTGTGCCACGTCGCTCAGG 1165
Db 22 GluAspGlnProTrpValLysArgLeuValMetValLysValValProThrCysLeuLys 41
QY 1166 GCCTTGTAGTAAATGCCCGGTAGGGGTGCTCTCTCTCGAGAAATCTGTGGACCTG 1225
Db 42 GluLeuLeuGluMetAlaArgGluGlyGlyAlaSerSerLeuLys---ThrValAspLeu 60
QY 1226 CACATTTCCAAAGCAGCCACCTCTCCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCAG 1285
Db 61 HisIleSerAsnSerGlnProIleSerLeuThrSerAspGlnTrpLysAlaCysLeuGln 80
QY 1286 GACTTGTGGAGGCGATGATTTCAGGGCCCTGGGGAGAGC 1327
Db 81 AspLeuValGluAspMetAspPheGlnAlaThrGlyGluThr 94

RESULT 5
US-10-264-049-2929
Sequence 2929, Application US/10264049
Publication No. US20040005579A1
GENERAL INFORMATION:
APPLICANT: Birse et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PA133PI
CURRENT APPLICATION NUMBER: US/10/264,049
CURRENT FILING DATE: 2002-10-04
PRIOR APPLICATION NUMBER: PCT/US01/18569
PRIOR FILING DATE: 2001-06-07
PRIOR APPLICATION NUMBER: US 60/209,467
PRIOR FILING DATE: 2000-06-07
NUMBER OF SEQ ID NOS: 4360
SOFTWARE: PatentIn Ver. 3.1
SEQ ID NO 2929
LENGTH: 395
TYPE: PRT
ORGANISM: Homo sapiens
US-10-264-049-2929

Alignment Scores:
Pred. No.: 1.32e-08 Length: 395
Score: 257.00 Matches: 70
Percent Similarity: 43.17% Conservative: 28

75 AspPheSerGlnProAspThrSerProAspThrAsnGlyGlyGlySerThrSerAspThr 94
QY 452 CAGGAAGACATCTGGATGAGTACTGGGTAAACATGTTGGTGGCCCACTCCAGATCCG 511
Db 95 GlnGluAspIleLeuAspGluLeuLeuGlyAsnMetValLeuAlaProLeuProAspPro 114
QY 512 GGACCCCAAGCTGGTGTAGCCCTCAGCCCTCAGCCCTCAGCCCTCAGCCCTCAGCCCTCAGC 571
Db 115 GlyProProSerLeuAlaValAlaProGluProCysProGlnProLeuLeuArgSerProSer 134
QY 572 TTGGACATCCACTCCCTTCCCAACCTGGGGCCCTCTGAGAACCCACTGAAGCGGCTG 631
Db 135 LeuAspAsnProThrProPheProAsnLeuGlyProSerGluAsnProLeuLysArgLeu 154
QY 632 TTGGTGGCGGGGAAGAGTGGAGTTCGAGGTTCAGAGCTTCTACCGGGCGGCCCAAGTC 691
Db 155 LeuValProGlyGluGluTrpGluPheGluValThrAlaPheTyrArgGlyArgGlnVal 174
QY 692 TTCAGAGACCATCTCTGCGCCGAGGCGCTGGTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 751
Db 175 PheGlnGlnThrIleSerCysProGluGlyLeuArgLeuValGlySerGluValGlyAsp 194
QY 752 AGGAGCTGCTGGATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 811
Db 195 ArgThrLeuProGlyTrpProValThrLeuProAspProGlyMetSerLeuThrAspArg 214
QY 812 GGAGTGATGAGCTAGCTGAGGCTGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 871
Db 215 GlyValMetSerTyrValArgHisValLeuSerCysLeuGlyGlyGlyLeuAlaLeuTrp 234
QY 872 CGGGCGGGCAGTGGCTTGGGCCCGCAGCGCTGGGGCTGGGCATGCCACATACCTAGTGGCAGTG 931
Db 235 ArgAlaGlyGlnTrpLeuTrpAlaGlnArgLeuGlyHisCysHisThrTyrTrpAlaVal 254
QY 932 ACGGAGGAGCTGCTCCCAACAGCGGCGCTGATGGGAGTCCCGAGGTCGCCAAGACAAAG 991
Db 255 SerGluGluLeuLeuProAsnSerGlyHisGlyProAspGlyGluValProLysAspLys 274
QY 992 GAAGGAGGGGTGTTGACCTGGGGCCCTTCATTGTA-----GATCTG 1033
Db 275 GluGlyGlyValPheAspLeuGlyProPheIleValGlySerLeuGlyProProAspLeu 294
QY 1034 ATTACCTTCAGGAAGAGGAGCGACCTCACACGCTATGCCCTCTGTTCTGTTGGGG 1093
Db 295 IleThrPheThrGluGlySerGlyArgSerProArgTyrAlaLeuTrpPheCysValGly 314
QY 1094 GAGTCATGGCCCGAGCAGCGCTGACCAAGAGGCTCGTGTAGTGTCAAGGTTGTGCC 1153
Db 315 GluSerTrpProGlnAspGlnProTrpThrLysArgLeuValMetValLysValValPro 334
QY 1154 ACGTGCCTCAGGCGCTTGTAGAAATGGCCCGGTAGGGGTGCTCTCTCTCTGAGAAAT 1213
Db 335 ThrCysLeuArgAlaLeuValGluMetAlaArgValGlyGlyAlaSerSerLeuGluAsn 354
QY 1214 ACTGTGGACCTGCACATTTCCACAGCCCACTCTCCCTCCTCCTCCTCCTCCTCCTCCTCCTC 1273
Db 355 ThrValAspLeuHisIleSerAsnSerHisProLeuSerLeuThrSerAspGlnTrpLys 374
QY 1274 GCCTACTGAGGACCTTGGTGGAGGCGATGATTTCAGGGCCCTGGGGAGAGC 1327
Db 375 AlaTyrLeuGlnAspLeuValGluGlyMetAspPheGlnGlyProGlyGluSer 392

RESULT 4
US-10-258-666-31
Sequence 31, Application US/10258665
Publication No. US20040005578A1
GENERAL INFORMATION:
APPLICANT: Yamada, Yoji
APPLICANT: Sekine, Susumu
APPLICANT: Kikuchi, Yasuhiro
APPLICANT: Sakurada, Kazuhiro
APPLICANT: Kyowa Hako Kogyo Co., Ltd.
TITLE OF INVENTION: Myocardial Cell Proliferation-Associated Genes

10508	ArgT	SerVal	SerSerAla	ArgThrCys	ProSerThrTrp	GlySerSerSerArg	Cys	10527
879	CGG	CCGCC	CCAGAG	CCAGT	-----	CCCCACCCAGG	AGCTCAGCATGCTCACGTAG	823
10528	Cys	Ala	ArgAla	TrpSerSerGly	ProArgPro	GlySerThrGly	Thr-----	10543
822	CTC	ATCATC	CCCTCTG	CTGTCTG	CAGGACATGC	CAGGCTGTGC	CAGTGTGACATGCCATCCA	763
10544	---Arg	SerPro	SerSerThrAla	ArgTrpSerSerThr	GlySerGly	ThrSerArg	Pro	10562
762	GGC	AGCGTCTGT	TCCAC	TTCGGAC	-----	-----	CC	733
10563	Arg	SerProAla	CysGlySerArg	ThrArgSerThr	GlyGlyGly	GluProAla	GluPr	10582
732	ACC	AGCC	-----	-----	CGAGCCCTC	CGGCAGG	ATGCTCTGC	697
10582	oPro	AlaAlaAla	GlyArgAla	ArgProGlyAla	GlyAlaPro	GlyArg	-----	10600
696	TGA	AGACTG	TCGGGC	-----	CCCGT	-----	AGAAG	670
10600	aHis	ArgProGly	GlyHisArg	ArgAlaLeuPro	GlyAlaAspArg	SerAlaPro	ArgAr	10620
669	GCT	GTCACTCG	AACTCC	CACTCTCC	CCCG	-----	GCACCAACAGCCGCTT	616
10620	gLeu	GlyLeuArg	ProSerAla	ArgProGln	ArgAlaGlyArg	AlaAlaArg	ArgGl	10640
615	T-----	-----	-----	-----	-----	-----	TCTCA	610
10640	yVal	GlyAlaLeu	AlaArgLeu	ArgSerGlyHis	LeuAlaValPro	ArgProAla	ProAr	10660
609	GAGG	CCCCAGG	TTTGGG	AAAGGAG	-----	TGGGATTC	CAAGTGGGCTCCG	556
10660	gArg	ProAlaGly	AspGlyArg	AspAlaArg	ArgArgGlyPro	AspGlySer	ArgAl	10680
555	GCT	GAGCGC	AGGCTCAG	-----	-----	GGGCTAC	GACC	526
10680	aGly	ArgGlyArg	AlaValArg	ProAlaGlyArg	LeuValArg	CysAlaLeu	Hispr	10700
525	AGC	TGGGGG	TCCGGATCT	GGAGTGGG	CGCCACAC	CATGTATCC	CACTATCC	466
10700	oGly	ProGlyThr	ArgAsp	-----	GlyProGlyPro	AlaSerAla	ArgGlyAlaPr	10717
465	AGA	ATGCTCT	CCTGGGTAT	CAGAAGTACT	GCCTCC	ACCATTGGTGT	CCGGAGAGTGTCT	406
10717	oAla	AlaAlaAla	GlyArgGln	AlaGlyAlaValHis	AlaTrp	-----	ArgLe	10733
405	GGT	TGG	-----	AAAAAGTCC	CAACTCTG	AGTTTCA	CAACTCGTATATTTATGT	355
10733	uGly	GlyAlaArg	ArgAlaGly	ProValAlaLeu	AspArgArg	Arg	-----	10750
354	GGT	CGTGAG	GGTCTTGTCT	CCGCTCTCTG	CTAAACGCA	ACCTTCTTT	CGGTTGAGG	295
10751	-----	GlyPro	GlyArgAla	GlyArgValHis	Arg	-----	G	10761
294	GC	AGAGCG	GAATCTCT	CTCCAGGTTGG	CAGGCTCG	CTTATCC	TATCCCGGGACATAT	235
10761	yPro	AsnGlyGln	AlaHisLeu	AlaArgThrAla	ArgSerAla	AlaPro	-----	10777
234	GC	ACCATGG	CTCGGCC	AGGCTCG	AGATTCCG	AAATCTCTCT	CTGTGTGCATCTGC	175
10778	-----	AlaLeu	ProArgPro	Gly	-----	-----	AlaArgLeuAl	10787
174	CGT	AGCGCTG	CTTCC	AAAGGATCG	GAAGCGCTG	CGGCTCTT	GTTCACCC	123
10787	aGly	CysArgSerGly	ArgAlaAlaAla	AlaAspPro	GlyArgPro	GlyArgPro	AlaAlaPr	10807
122	-----	-----	-----	AGGC	CAGCCCTC	CAAGTGC	CC	100
10807	oAla	ProArgArg	ArgLeuPro	ArgProGlyLeu	ArgArgPro	ArgArg	GlyPr	10827
99	AGT	CCAGCT	CGCACCA	-----	-----	GCAGG	CAGGAGNC	67
10827	oGly	ArgAlaValLeu	LeuProGlyHis	ProAlaAspArg	GlnGlnAlaAla	ArgAlaGly	Gl	10847

QY 66 CGTGGCTTTGGGG-----TTCCC 49
Db 10847 uAlaGlyGluGlyArgAspMetThrGlnProProSerProThrMetCysValLeuPr 10867
QY 48 ATGG 45
Db 10867 oTTP 10868
RESULT 7
US-09-809-391-374
; Sequence 374, Application US/09809391
; Publication No. US20030049618A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P2
; CURRENT APPLICATION NUMBER: US/09/809,391
; PRIOR FILING DATE: 2001-03-16
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 761
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 374
; LENGTH: 373
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (175)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (373)
; OTHER INFORMATION: Xaa equals stop translation
US-09-809-391-374
Alignment Scores:
Pred. No.: 6.93e-08 Length: 373
Score: 246.00 Matches: 104
Percent Similarity: 42.32% Conservative: 64
Best Local Similarity: 26.20% Mismatches: 161
Query Match: 9.11% Indels: 68
DB: Gaps: 20
US-09-975-253-1 (1-1426) x US-09-809-391-374 (1-373)
QY 329 GACCGGAGCAAGAC---CCTCACGACCACCAATAAAATCTACGAGTTTGTGAACTCAGGA 385
Db 3 AspGlyThrLysGluValProMetAsnProValLysIleTyrGln----- 17
QY 386 GTTGGGGACTTTTCCAGCCAGCACCTCT-----CCGACACCAATGGTGGAGGC 436
Db 18 ValCysAspIleProGlnProGlnGlySerIleLeAsnProGlySerThrGlySerAla 37
QY 437 AGTACTTCTGAT-----ACCCAGGAAGACATTTCTGGATCAGTTACTGGGT 481
Db 38 ProTrpAspGlyLysAspAsnAspValAspGluAspGluAspGluLeuAspGln 57
QY 482 AACATGGTGTGGCCCCACTCCACGATCCGGGACCCCGAGCTGGCTGGTACCCCTGAG 541
Db 58 SerGlnHisHisValProIleGlnAspThrPhe---ProPheLeuAsnIleAsnGlySer 76
QY 542 CCC---TGCCCTCAGCCCTCGGGAGCCCGAGCTGGACAATCCCACTCCC----- 589
Db 77 ProMetAlaProAlaSerValGlyAsnCysSerValGlyAsnCysSerProGluAlaVal 96
QY 590 TTCCCAAAACCTGGGGCCCTCTGAG-----AACCCACTGAAGCGGCTGTG 634
Db 97 TrpProLysThrGluProLeuGluMetGluValProGlnAlaProIleGlnProPheTyr 116
QY 635 GTCCCGGGGAAGAGTGG-----GAGTTCGAGGTGACAGCCTTC 673
Db 117 SerSerProGluLeuTrpIleSerSerLeuProMetThrAspLeuAspIleLysPheGln 136

1	PRIOR FILING DATE: 1997-05-23	
2	PRIOR APPLICATION NUMBER: 60/043,578	
3	PRIOR FILING DATE: 1997-04-11	
4	PRIOR APPLICATION NUMBER: 60/043,576	
5	PRIOR FILING DATE: 1997-04-11	
6	PRIOR APPLICATION NUMBER: 60/047,501	
7	PRIOR FILING DATE: 1997-05-23	
8	PRIOR APPLICATION NUMBER: 60/043,670	
9	PRIOR FILING DATE: 1997-04-11	
10	PRIOR APPLICATION NUMBER: 60/056,632	
11	PRIOR FILING DATE: 1997-08-22	
12	PRIOR APPLICATION NUMBER: 60/056,664	
13	PRIOR FILING DATE: 1997-08-22	
14	PRIOR APPLICATION NUMBER: 60/056,876	
15	PRIOR FILING DATE: 1997-08-22	
16	PRIOR APPLICATION NUMBER: 60/056,881	
17	PRIOR FILING DATE: 1997-08-22	
18	PRIOR APPLICATION NUMBER: 60/056,909	
19	PRIOR FILING DATE: 1997-08-22	
20	PRIOR APPLICATION NUMBER: 60/056,875	
21	PRIOR FILING DATE: 1997-08-22	
22	PRIOR APPLICATION NUMBER: 60/056,862	
23	PRIOR FILING DATE: 1997-08-22	
24	PRIOR APPLICATION NUMBER: 60/056,887	
25	PRIOR FILING DATE: 1997-08-22	
26	PRIOR APPLICATION NUMBER: 60/056,908	
27	PRIOR FILING DATE: 1997-08-22	
28	PRIOR APPLICATION NUMBER: 60/048,964	
29	PRIOR FILING DATE: 1997-06-06	
30	PRIOR APPLICATION NUMBER: 60/057,650	
31	PRIOR FILING DATE: 1997-09-05	
32	PRIOR APPLICATION NUMBER: 60/056,884	
33	PRIOR FILING DATE: 1997-08-22	
34	PRIOR APPLICATION NUMBER: 60/057,669	
35	PRIOR FILING DATE: 1997-09-05	

Alignment Scores:	6.93e-08	Length:	373
Pred. No.:	246.00	Matches:	104
Score:	42.32%	Conservative:	64
Percent Similarity:	42.32%	Mismatches:	161
Best Local Similarity:	26.20%	Indels:	68
Query Match:	9.11%	Gaps:	20
DB:	10		

[illegible]

```

QY 674 TACCGGGCCGCAAGTCTTCAG---CAGACCATCTCTGCGCGGAGGCGCTGGCGTG 730
Db 137 TyrArgGlyLysGluThrMetThrValSerAsnProGlnGlyCysArgLeu 156
QY 731 GTGGGTCGCAAGTGGGA-----GACAGGACGCTGCTGGATGGCCAGTCACACTG 781
Db 157 PheTyrGlyAspLeuGlyProMetProAspGlnGluLeuPheGlyProVal***Leu 176
QY 782 CCAGAC-----CCTGGCATGTC---CTGACAGACAGGGAGTGATGAGTACAGTG 829
Db 177 GluGlnValPheProGlyProGluHisIleThrAsnGluLysGlnLysLeuPheThr 196
QY 830 AGGCATGTCTAGCTGCTGGTGGGGAGCTCTCTCTGGCGGCGCGGAGTGGCTC 889
Db 197 SerLysLeuLeuAspValMetAspArgGlyLeuIleLeuGluValSerGlyHisAlaIle 216
QY 890 TGGGCCACGCGCTGGGCACCTGCCACACATACTGGGCA-----GTGAGCGAGGAG 940
Db 217 TyrAlaIleArgLeuCysGlnCysLysValTyrTrpSerGlyProCysAlaProSerLeu 236
QY 941 CTGCTCCCAACAGCGGGCATGGGCTGATGGAGGTCCCAAGACAAAGGAAGAGGC 1000
Db 237 ValAlaProAsnLeu-----IleGluArgGlnLysLysValLys 249
QY 1001 GTGTTTGACCTGGGCGCCCTTCATTGTAGATCTGATTACCTTCACGGAAGAGC---GGA 1057
Db 250 LeuPheCysLeuGluThrPheLeuSerAspLeuIleAlaHisGlnLysGlyGlnLeu 269
QY 1058 CGCTCACCAGCGTATGCCCTCTGTTCTGTGGGGAGTCTATGGCCCCCAGGACGCG 1117
Db 270 LysGlnProProGluIleTyrLeuCysPheGlyGluGluTyrProAspGlyLysPro 289
QY 1118 TGGACCAAGAGCTGCTGATGTGTAAGTTGTGGCCACGTCGCTCAGGSCCTTGTAGAA 1177
Db 290 LeuGluArgLysLeuIleLeuValGlnValIleProValValAlaArgMetIleTyrGlu 309
QY 1178 ATGGCCCGGGTAGGGGGTCCCTCCCTCCCTGGAGAT---ACTGTGAGCTGCACATTTCC 1234
Db 310 MetPheSerGlyAspPheThrArgSerPheSerGlySerValArgLeuGlnIleSer 329
QY 1235 AACAGCCACCCACTCTCCCTCACCTCCGACCAAGCTACAAGGCGCTACTGCGAGCTTGTG 1294
Db 330 ThrProAsp-----IleLysAspAsnIleValAlaGlnLeuLysGlnLeuTyr 345
QY 1295 GAGGCATGATTCCAGGGCCCTGGGAGAGACTGAGCCCTGCTCCTCATGTGTGCT 1354
Db 346 ArgIleLeuGlnThrGlnGluSerTrpGln----- 355
QY 1355 CCAACCCCGCTGTTCCCAACACCTCAACCAATAAAGTGTCTCTGCTATG 1405
Db 356 -----ProMetGlnProThrProSerMet---GlnLeuProProAlaLeu 369

RESULT 9
US-10-164-861-374
; Sequence 374, Application US/10164861
; Publication No. US20030225248A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: PZ002P1
; CURRENT APPLICATION NUMBER: US/10/164,861
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: US/09/149,476
; PRIOR FILING DATE: 1998-09-08
; PRIOR APPLICATION NUMBER: PCT/US98/04493
; PRIOR FILING DATE: 1998-03-06
; NUMBER OF SEQ ID NOS: 757
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 374
; LENGTH: 373
; TYPE: PRT
; ORGANISM: Homo sapiens

```

```

; FEATURE:
; NAME/KEY: SITE
; LOCATION: (175)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (373)
; OTHER INFORMATION: Xaa equals stop translation
US-10-164-861-374

```

```

Alignment Scores:
Pred. No.: 6,93e-08 Length: 373
Score: 246.00 Matches: 104
Percent Similarity: 42.32% Conservative: 64
Best Local Similarity: 26.20% Mismatches: 161
Query Match: 9.11% Indels: 68
DB: 14 Gaps: 20

US-09-975-253-1 (1-1426) x US-10-164-861-374 (1-373)
QY 329 GACCGAGCAAGGAC---CCTCAGCAGCCACATAAAATCTACGAGTTTGTGAATCAGGA 385
Db 3 AspGlyThrLysGluValProMetAsnProValLysIleTyrGln----- 17
QY 386 GTTGGGACTTTTCCAGCCAGACACCTCT-----CCGAGCACCAATCGTGGAGGC 436
Db 18 ValCysAspIleProGlnProGlnGlySerIleIleAsnProGlySerThrGlySerAla 37
QY 437 AGTACTTCTGAT-----ACCCAGGAGAGACATCTCTGATGAGTACTGGGT 481
Db 38 ProTrpAspGluLysAspAsnAspValAspGluGluAspGluLysLeuAspGln 57
QY 482 AACATGTTGTGGCCCCACTCCAGATCCGGGACCCCAAGCCTGGCTGATGCCCTGAG 541
Db 58 SerGlnHisHisValProIleGlnAspThrPhe---PropheLeuAsnIleAsnGlySer 76
QY 542 CCC---TGCCCTCAGCCCTCGGAGCCCGACCTTGGACAATCCCACTCCC----- 589
Db 77 ProMetAlaProAlaSerValGlyAsnCysSerValGlyAsnCysSerProGluAlaVal 96
QY 590 TTCCCAACACCTGGGCGCCCTCTGAG-----AACCCACTGAAGCGGCTGTG 634
Db 97 TrpProLysThrGluProLeuGluMetGluValProGlnAlaProIleGlnProPheTyr 116
QY 635 GTGCGCGGGGAAGAGTGG---GAGTTTCGAGGTGACAGCCCTC 673
Db 117 SerSerProGluLeuTyrTrpIleSerSerLeuProMetThrAspLeuAspIleLysPheGln 136
QY 674 TACCGGGCCGCAAGTCTTCAG---CAGACCATCTCTGCGCGGAGGCGCTGGCGTG 730
Db 137 TyrArgGlyLysGluTyrGlyGlnThrMetThrValSerAsnProGlnGlyCysArgLeu 156
QY 731 GTGGGTCCGAAGTGGGA-----GACAGGACGCTGCTGGATGGCCAGTCACACTG 781
Db 157 PheTyrGlyAspLeuGlyProMetProAspGlnGluLeuPheGlyProVal***Leu 176
QY 782 CCAGAC-----CCTGGCATGTC---CTGACAGACAGGGAGTGCATGAGTACAGTG 829
Db 177 GluGlnValLysPheProGlyProGluHisIleThrAsnGluLysGlnLysLeuPheThr 196
QY 830 AGGCATGTCTAGCTGCTGGTGGGGAGCTGCTCTCTGGCGGCGCGGAGTGGCTC 889
Db 197 SerLysLeuLeuAspValMetAspArgGlyLeuIleLeuGluValSerGlyHisAlaIle 216
QY 890 TGGGCCACGCGCTGGGCACCTGCCACACATACTGGGCA-----GTGAGCGAGGAG 940
Db 217 TyrAlaIleArgLeuCysGlnCysLysValTyrTrpSerGlyProCysAlaProSerLeu 236
QY 941 CTGCTCCCAACAGCGGGCATGGGCGCTGATGGAGGTCCCAAGACAAAGGAAGAGGC 1000
Db 237 ValAlaProAsnLeu-----IleGluArgGlnLysLysValLys 249
QY 1001 GTGTTTGACCTGGGCGCCCTTCATTGTAGATCTGATTACCTTCACGGAAGAGC---GGA 1057

```



```

Db      2539 CysAlaGlyCysProAlaThrLeuThrIleGlyArg-
QY      371 ACTCGTAGATTTATGTGGTCTGTAGGTCCTTGTCTCGGTCTCTGCTAAACGCAACC 312
Db      2554 ThrCysProLeuTyrValGlySerArgGlyAlaLeu-
QY      311 CTCTTTGGTGGTTAGGCGAGCGGAAATCTCTTCCAGGTGGCAGGTCTGGCTTAT 252
Db      2566 -----GlyGlnAlaAspArgLeuLysAlaGluIleGlyArgLeuAenGly 2580
QY      251 CCTCCCGGAAACATATGCACCATGTGGCTCGCCAGCGCTGGAAGATTCGGAATCCT 192
Db      2581 ProValArg-----SerTTPProAlaProAlaValProAlaThrArgArgSer 2596
QY      191 CTGTCTGTCTATCTCCCTA-----GGCGGTGTTCCAGGGATGCGGAAGCGGTGC 138
Db      2597 ArgAlaAlaSerProThrThrGlyProGlyArgAlaGlyArgAspArgGlyProGlyArg 2616
QY      137 GGCTCTGTGTACCCAGGCGACGCTCCAGTTCAGTTCGCCAGGTCCAGCTCGGACACGAGCC 78
Db      2617 GlyAlaArgSerValProSerArgProProVal-----ProAlaTTPArgArgCys 2633
QY      77 AGGGCA 72
Db      2634 ArgAla 2635

RESULT 11
US-09-805-020-46
; Sequence 46, Application US/09805020
; Publication No. US20020086384A1
; GENERAL INFORMATION:
; APPLICANT: LEVINE, Zurit
; TITLE OF INVENTION: SPLICE VARIANTS OF ONCOGENES
; FILE REFERENCE: 2786-0168P
; CURRENT APPLICATION NUMBER: US/09/805,020
; CURRENT FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 46
; LENGTH: 292
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-805-020-46

Alignment Scores:
Pred. No.:      1.67e-07      Length:      292
Score:          240.00      Matches:      65
Percent Similarity: 43.66%      Conservative: 28
Best Local Similarity: 30.52%      Mismatches: 85
Query Match:      9          Indels:      35
DB:              9          Gaps:       6

US-09-975-253-1 (1-1426) x US-09-805-020-46 (1-292)
QY      47 ATGGGAACCCAAAGCCAGGTCCTCGCTCGGTGTGTGTCAGCTGGAGCTGGGCAA 106
Db      1 MetProIleThrArgMetArgMetArgProIlePueGluMetGlnIleAsnSerAsnGln 20
QY      107 CTGAGGCGGTGGCTGGTGAACAAGACGCGCTTCCGCTTCCGCTATCCCTTGAAGCAC 166
Db      21 IleProGlyLeuIleTTPileAsnLysGluGluMetIlePheGlnIleProTTPylsHis 40
QY      167 GGCCTACGCGAGGTGCACAG-----CAGGAGATTTCCGAATCTTCCAGGCTGGGCCGAG 223
Db      41 AlaAlaLysHisGlyTTPAspIleAsnLysAspAlaCysLeuPheArgSerTTPAlaIle 60
QY      224 GCCACTGTGTGCATATGTTCCCGGAGGATAGCCAGACCTGCCAACCTGGAAGAGGAT 283
Db      61 HisThrGlyArgTyrLysAlaGlyGluLysGluProAspProLysThrTTPylsAlaAsn 80
QY      284 TTCGCTCTGCCTCAACCGCAAAAGAGGTTGGGTTTATGAGAGGACCGGACGCAAGGAC 343

```

```

Db      81 PheArgCysAlaMetAsnSerLeuProAspIleGluGluValLysAspGlnSerArgAsn 100
QY      344 CCT---CAGCACCACATAAATCTACGAGTTTGTG-----
Db      101 LysGlySerSerAlaValArgValTyrArgMetLeuProLeuThrLysAsnGlnArg 120
QY      377 -----NACTCAGAGATTGGGACATTTTCCAG 403
Db      121 LysGluArgLysSerLysSerArgAspAlaLysSerLysAlaLysArgLysSerCys 140
QY      404 CCAGACACCTCTCCGAGACCAATGTGTGAGCAGTACTTCTCATACCCAGGAAGACATT 463
Db      141 GlyAspSerSerProAspThrPheSerAspGlyLeuSerSerThrLeuProAsp-As 160
QY      464 CTGGATGATTTACTGTGTAACATGTGTGTGGCCCACTCCAGATCCGCGGACCCCAAGC 523
Db      160 phisSerSerTyrThrValProGlyTyrMetGlnAspLeuGluValGln----- 177
QY      524 CTGGCTGTAGCCCTGAGCCCTGAGCCCTCAGCCCTCGCGGAGCC-----CCAGC 571
Db      178 -----AlaLeuThrProAlaLeuSerProCysAlaValSerSerThrLeuProAs 194
QY      572 TTGGACAATCCCA-----CTCCTTCCCAAAACC 599
Db      194 pTTPhisIleProValGluValProAspSerThr 206

RESULT 12
US-10-084-846A-6
; Sequence 6, Application US/10084846A
; Publication No. US20040006026A1
; GENERAL INFORMATION:
; APPLICANT: WEITNAUER, GABRIELE
; APPLICANT: MUHLENWEG, AGNES
; APPLICANT: TREPZER, AXEL
; APPLICANT: BECHTHOLD, ANDREAS
; TITLE OF INVENTION: AVILAMYCIN DERIVATIVES
; FILE REFERENCE: 1974-005
; CURRENT APPLICATION NUMBER: US/10/084,846A
; CURRENT FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: PCT/EP01/09815
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: DE 101 09 166.4
; PRIOR FILING DATE: 2001-02-25
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 6
; LENGTH: 19662
; TYPE: PRT
; ORGANISM: Streptomyces viridochromogenes
; FEATURE:
; OTHER INFORMATION: Protein 1: amino acid sequence encoded by coding strand 2.
; OTHER INFORMATION: Start codon: gga, Start position: nucleotide 1.
US-10-084-846A-6

Alignment Scores:
Pred. No.:      3.2e-07      Length:      19662
Score:          239.50      Matches:      157
Percent Similarity: 32.45%      Conservative: 27
Best Local Similarity: 27.69%      Mismatches: 209
Query Match:      8.95%      Indels:      174
DB:              15       Gaps:       33

US-09-975-253-1 (1-1426) x US-10-084-846A-6 (1-19662)
QY      1379 AGTGGTGG-----GGAACAGGGGGTTGAG 1353
Db      12610 ArgTTPTrpAlaAlaSerArgCysTTPProAlaArgProGlyArgGly----- 12626
QY      1352 GCACACCATGAGGAGCGAGGCTCAGCTCTCCCGAGGCCCTGGAATCCATGCCCTCCA 1293
Db      12627 -----AlaSerProAlaGlyGlyProProSerSer 12638
QY      1292 CCAAGTCTCGAGTAGGCTTGTACTGTGTCGGAGGTGAGGGAGTGGGTGTGTGG 1233

```

```

12961 -----GlySerArgArgArg-----GlyArgAlaArgProLeu 12971
335 TCCGGTCTCTGCTAAACGCAACCTCTTTCGCGTTGAGGCGAGAGCGAAATTCCTCT 276
12972 AlaProArgArgTrpAsnAlaThr-----AlaGlyGlyAlaAlaArg 12985
275 TCCAGGTGGCAGCTCTG-----GCTTATCCCTCC 246
12986 GlyArgAlaAlaGlyLeuArgGlySerArgSerAlaGlyAlaProThrSerProArgSer 13005
245 CGGAAACATATGCACCATGCGCTCGG-----CCAGGCGCTCGA--- 207
13006 ArgArgGlySerGlyArgArgProArgArgArgGlyArgSerGlyProArgAlaGlyCys 13025
206 -----AGATTCCGAAATCTCTCTGTCATCTCGCTAGGCGCT 165
13026 AlaValAlaSerValArgArgThrArgProArgValArgProAlaSerArgGly 13045
164 GCTTCC-----AAGGATGCGGAAGCGGCTCGCTCTTGTCCACCGGCACGC 114
13046 SerSerAlaArgProArgGlyAlaGlyThrAlaArgArgThrProGlyGlyArgProArg 13065
113 CCTCCA-----GTTCCCGCAGTCCAGCTGCGACACACCGCCAGGG--- 74
13066 ProProArgSerArgArgArgAlaAlaProGlyCysAlaGlySerPro-ProSerThrAr 13085
73 -----CAGGANCCGCTGCTTGGGGTCCCATGTCGCGCTACGATGGAAGTCCGG 21
13095 gArgArgArgSerProTrpProAlaGlyProArgAlaProSerCys-----ArgAlaAl 13103
20 GCGTCGCGGCGAGCTGGAAC 2
13103 aCysAlaValArgTrpAsp 13109

RESULT 13
US-10-084-846A-8
; Sequence 8, Application US/10084846A
; Publication No. US20040006026A1
; GENERAL INFORMATION:
; APPLICANT: WEITNAUER, GABRIELE
; APPLICANT: MUHLNBERG, AGNES
; APPLICANT: TREFFER, AXEL
; APPLICANT: BECHTHOLD, ANDREAS
; TITLE OF INVENTION: AVILAMYCIN DERIVATIVES
; FILE REFERENCE: 1974-005
; CURRENT APPLICATION NUMBER: US/10/084,846A
; CURRENT FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: PCT/EP01/09815
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: DE 101 09 166.4
; PRIOR FILING DATE: 2001-02-25
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: Patent in Ver. 3.2
; SEQ ID NO 8
; LENGTH: 19608
; TYPE: PRT
; ORGANISM: Streptomyces viridochromogenes
; FEATURE:
; OTHER INFORMATION: Protein 3: amino acid sequence encoded by coding strand 1.
; OTHER INFORMATION: Start codon: atc, Start position: nucleotide 3.
US-10-084-846A-8

Alignment Scores:
Pred. No.: 6,83e-07 Length: 19608
Score: 234.50 Matches: 153
Percent Similarity: 33.16% Conservative: 39
Best Local Similarity: 26.42% Mismatches: 208
Query Match: 8.78% Indels: 179
DB: 15 Gaps: 32

US-09-975-253-1 (1-1426) x US-10-084-846A-8 (1-19608)
QY 1379 AGGTGGTGGGAACAGGGGGTTCGAGGCACACCATGAGGAGCGAGGCTCAGCTCTCC 1320

```

Db 11090 ArgArgTrpAlaArgProGly---ArgProArgThrArgArgArgAlaAlaProValPro 11108
 QY 1319 CAGGGCCCTGGAATCCATGCTCCACCAAGTCTCCAGGTAGCCCTTGTACTGTGTCGG 1260
 Db 11109 ArgGlyAlaAlaGlyProValProGlyProSerProAlaSerProSerCysAlaArgGly 11128
 QY 1259 AGGTGAGGAGAGA---GTGGGTGGCTTTGGAATGTGA 1224
 Db 11129 Arg---GlyArgAspAlaHisArgProArgGlySerSerSerArgArgSerAla 11147
 QY 1223 GTTCCACAGTATTCT---CCAGGGAGG 1200
 Db 11148 GlyProArgArgSerGlyAlaValArgArgProProAspGlyArgSerArgProGlyArg 11167
 QY 1199 AGGACCCCTACCCGGGCA---TTTCTACCAAGCCCTGAGGACAGTGGGCACACACT 1143
 Db 11168 Glu---ProGlyProArgSerArgSerArgProValGlyThrAlaAlaArgSerPro 11184
 QY 1142 TGACCATCAGAGCCTCT---TGTCCACGCTGGTCTCTGGGCCATGACTCCCCACAC 1086
 Db 11185 ArgProGlyProValSerProTrpSerTrp---MetArgProProPhe 11199
 QY 1085 AGAACAGAGGCGATAGCGTGGTGGAGCTCCGCTTCTCCGTGAAGGTAATCAGATCTA 1026
 Db 11200 AlaserSerGlyGln---ArgHisGluArgGlyThrProAlaLeuArg---GlnAla 11217
 QY 1025 CAAATGAAGGCCCCAGCGTCAACACAGCCTCTCTTCTGTCTTGGGACCTCCCATCAG 966
 Db 11217 laThrMetSerGlnThrSerAlaValAlaArgSerCysSerThrSerValysP 11237
 QY 965 GCCATGCGCGTGTGGGAGAGCTCTCGCTCACTCCCTCCAGTATGTG--- 916
 Db 11237 roAlaCysSerArgAlaAlaArgSerAspGlyArgSerSerGlnArgThrSerThrLeuA 11257
 QY 915 ---TGGCAGTGCCTCCAGCCGCTGG---GCCAGAGCCACT 882
 Db 11257 rgProArgAlaTrpGlySerGlyTrpSerSerArgThrGlyTrpAlaAlaTyArgArg- 11276
 QY 881 GCCCGCCCGCAGAGAGCAGTCCCCACCCAGGAGCTCAGCATCCCTCAGCTAGC 822
 Db 11277 GlyArgSerPheArgProSerCysProSerSerGlyTrpArgGlyArgGlySerArg 11296
 QY 821 TCATCACTCCCTGTCTGTCAGGACATCCAGGCTCTCCAGTGTGAGTGGCCATCCAG 762
 Db 11297 ArgSer---ArgCysArgA-gAlaAlaProArgAlaAlaAlaArg 11310
 QY 761 GCAGCGTCTGTCTC---CCACTCGGACCCCA 729
 Db 11311 GlyArgGlyAlaLeuArgAlaTrpArgArgSerGlyArgAlaArgArgThrProSer 11330
 QY 728 GCC---GCAGG---CCTCCGGCAGGAGATGG 702
 Db 11331 ThrProArgArgAlaGlyValProArgTrpArgAlaProProArgThrArgArg 11350
 QY 701 TCTGTGGAAGACTTGGGGCC---GGTAGAGGCTGTCACTCGA 657
 Db 11351 GlyAlaGlySerLeuArgGlyProProGlyArgGlySerArgArgAlaValProThr 11370
 QY 656 ACTCCACTCTTCCCGCCGACCAACAGCCGCTTCCAGTGGTCTCAGAGGCGCCCGAGT 597
 Db 11371 SerProGlyCysProArgProPro---AlaArgAlaPro--- 11383
 QY 596 TTGGAGGGAGTGGATTGTCAGAGTGGGCTCCGAGGGGTGAGGCGAGGCTCAG 537
 Db 11384 ---GlyArgAlaTrp---Pro-ProGlyThrProArgThrArgArgProAlaPr 11400
 QY 536 GGGCTACAGCAGGCTTGGGGTCCCGGATCTGG---CAGTGG--- 497
 Db 11400 oGlyAlaAlaGlySerAlaGlySerArgAlaTrpArgArgTrpArgThrGlyProArgC 11420
 QY 496 ---GGCCAAACACATGTT---ACCCA 477

Db 11420 sThrGlnSerArgSerProThrAlaArgValHisValProAspAlaGluSerGlnAspAr 11440
 QY 476 GTAACCTCATCAGAAATGCTTCTCGGTATCAGAGTACTGCCTCCACCATGTGTGTCGG 417
 Db 11440 sSerLeuLeuProTyrlzHisCysGlnValArgPheGlnAlaArgArgLeuGlyLeu-G 11460
 QY 416 GAGAGGTGTCTGCTGGGAAAAGTCC---CCAACTCCTCAGT 378
 Db 11460 ly---lleThrGlyArgArgSerAlaGlyArgProHisThrAlaProThrPro--- 11477
 QY 377 TCACAAACTCGTAGATTTTATGTTGGTCTGTGAGGCTCTTCTCCCGTCTCTCTCTAAAC 318
 Db 11478 ---GlyArgAlaLeuArgArgSerArgSerA 11487
 QY 317 GCAACCTTCTTTCGCTTTCAGGCGAGAGCGGAAATCTCTTCCAGGTGGCAGGTCTG 258
 Db 11487 rgArgSerAlaValArgGlyArgArg-ArgProThrArgArgSerGlyArg 11506
 QY 257 GCTTAT---CCTCCCGGGAACATATGCAC---CAGTGG 225
 Db 11507 AlaArgProAlaGlyCysProProSerArgArgArgProArgGlyProProArgTrp 11526
 QY 224 ---CCTCGGCCGCTCTGGAAGATTC--- 201
 Db 11527 ArgCysArgArgProSerProGlyArgAlaGlyAlaArgTrpGlyGlyTrpGly 11546
 QY 200 CGAAATCTCTGCTGTGCATCTGCTAGGCGGTCTTCCAGGGATCGGAGCGCG 141
 Db 11547 HisArgProHisAlaArgProGlyProAlaProAlaArgCysTrpCysGlyThrSer 11566
 QY 140 TCGGCTCTTGTTCACCCAGGCGCCCTCCAGTTGCCCCAGGT--- 96
 Db 11567 ArgArgArgProGlyProArgProSerProSerArgGlyArgAlaArgArgProArgCys 11586
 QY 95 CCAGTGCACACACAGC---AGGGCA 72
 Db 11587 ProArgSerThrProAlaProAspArgArgArgArgProArgSerAlaHisArgAla 11606
 QY 71 GGAGCCGTGCTTGGGGTTCCTCCAGTTCCTCCAGTTCAGTGAAG 27
 Db 11607 Gly---SerGlyArgGlyArg 11613

RESULT 14
 US-10-755-889-356
 ; Sequence 356, Application US/10755889
 ; Publication No. US20040171823A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bristol-Myers Squibb Company
 ; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-KB
 ; TITLE OF INVENTION: PATHWAY
 ; FILE REFERENCE: D0284 NP
 ; CURRENT APPLICATION NUMBER: US/10/755,889
 ; CURRENT FILING DATE: 2004-01-13
 ; PRIOR APPLICATION NUMBER: U.S. 60/440,068
 ; PRIOR FILING DATE: 2003-01-14
 ; PRIOR APPLICATION NUMBER: U.S. 60/469,757
 ; PRIOR FILING DATE: 2003-05-12
 ; NUMBER OF SEQ ID NOS: 823
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 356
 ; LENGTH: 164
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-755-889-356

Alignment Scores:
 Pred. No.: 5,59e-07 Length: 164
 Score: 231.50 Matches: 47
 Percent Similarity: 55.75% Conservative: 16
 Best Local Similarity: 41.59% Mismatches: 39
 Query Match: 16 Indels: 11
 DB: 3 Gaps: 3

Search completed: October 28, 2004, 15:50:49
Job time : 259 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: October 28, 2004, 14:56:31 ; Search time 296.5 Seconds

(without alignments)
5534.464 Million cell updates/sec

Title: US-09-975-253-1

Perfect score: 2699

Sequence: 1 ggttcacgtgccgcagc.....aaaaaaaaaaaaaaaaaaaaa 1426

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 3650362

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-Q/cgn2.1/USPTO spoop/US09975253/runat 28102004 111509 19624/app query.fasta 1.1607
-DB=Uniprot_02 -QFMT=fastan -SUFFIX=n2p.rup -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS-human40 cdi -LIST=45
-DOCALIGN=200 -THR SCORE=0 -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=8to -NORM=EXT -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USERS=US09975253@cgn 1 1 295@runat 28102004 111509 19624 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Uniprot_02.*

1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2458.5	91.1	452	Q7Z5G6	Q7Z5G6 homo sapien
2	2455.5	91.0	452	Q96GL3	Q96GL3 homo sapien
3	2326	86.2	427	1 IRF3 HUMAN	Q14653 homo sapien
4	2326	86.2	427	AAH71721	AAH71721 homo sapi
5	1742	64.5	419	Q764M6	Q764M6 sus scrofa
6	1742	64.5	419	2 BAD06317	BAD06317 sus scrof
7	1553	57.5	419	1 IRF3 MOUSE	Q7Z5G6
8	630	23.3	466	2 Q6RCM8	Q6RCM8 homo sapien
9	504.5	18.7	488	2 AAR90325	Aar90325 homo sapi
10	504.5	18.7	488	2 AAR90326	Aar90326 homo sapi
11	504.5	18.7	488	2 AAR90326	P56477 mus musculu
12	503.5	18.7	497	1 IRF5 MOUSE	Q6RCM8
13	493.5	18.5	498	1 IRF5 HUMAN	Aar90326 homo sapi
14	460.5	17.1	492	2 Q6PGZ7	Q6PGZ7 brachydanio
15	460.5	17.1	492	2 AAH56772	AAH56772 brachydanio
16	460	17.0	460	2 Q6GR23	Q6GR23 xenopus lae

17	459.5	17.0	491	2	Q90ZD4	Q90ZD4 gallus gall
18	457.5	17.0	491	1	IRF3 CHICK	Q90ZD4 gallus gall
19	451.5	16.7	467	1	IRF6 HUMAN	Q14896 homo sapien
20	449.5	16.7	467	1	Q8WNO4	Q8WNO4 sus scrofa
21	448	16.6	503	1	IRF7 HUMAN	Q23985 homo sapien
22	447	16.6	460	2	Q6DDZ6	Q6DDZ6 xenopus lae
23	439	16.3	467	1	IRF6 MOUSE	P97431 mus musculu
24	438.5	16.2	459	2	Q57578	O57578 xenopus lae
25	437.5	16.2	467	2	Q9N136	Q9N136 ovis aries
26	436	16.2	457	1	IRF7 MOUSE	P70434 mus musculu
27	436	16.2	457	1	BAC37723	BAC37723 mus musculu
28	431	16.0	467	2	Q91VD0	Q91VD0 mus musculu
29	400.5	14.8	421	2	Q802H9	Q802H9 carassius a
30	400	14.8	392	2	Q6NY63	Q6NY63 brachydanio
31	400	14.8	392	2	AAH66724	AAH66724 brachydanio
32	390.5	14.5	451	1	IRF4 HUMAN	Q15306 homo sapien
33	389.5	14.4	449	2	BAC40840	BAC40840 mus muscu
34	388	14.4	450	1	IRF4 MOUSE	Q64287 mus musculu
35	388	14.4	450	2	BAC40969	BAC40969 mus muscu
36	383.5	14.2	409	2	Q98TX6	Q98TX6 gallus gall
37	378.5	14.0	440	2	Q99419	Q99419 homo sapien
38	368.5	13.7	425	1	ICSB CHICK	Q90871 gallus gall
39	366.5	13.6	445	2	Q98TX7	Q98TX7 gallus gall
40	365.5	13.5	83	2	O54726	O54726 mus musculu
41	361.5	13.4	426	1	ICSB_HUMAN	Q02556 homo sapien
42	359	13.3	423	2	Q6PE40	Q6PE40 brachydanio
43	359	13.3	423	2	AAH58298	AAH58298 brachydanio
44	349.5	12.9	416	2	Q90W10	Q90W10 gallus gall
45	343.5	12.7	424	1	ICSB_MOUSE	P23611 mus musculu

ALIGNMENTS

RESULT 1		PRELIMINARY;		PRT;	452 AA.
Q7Z5G6					
ID	Q7Z5G6				
AC	Q7Z5G6				
DT	01-OCT-2003 (Tremblrel. 25, Created)				
DT	01-OCT-2003 (Tremblrel. 25, Last sequence update)				
DT	01-MAR-2004 (Tremblrel. 26, Last annotation update)				
DE	IRF3 protein.				
OS	Homo sapiens (Human)				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Kidney;				
RX	MEDLINE=2386257; PubMed=12477932;				
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,				
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,				
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,				
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,				
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,				
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,				
RA	Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,				
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,				
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,				
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay J.J., Hulyk S.W.,				
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gilts R.A.,				
RA	Faney J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,				
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.C.,				
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,				
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.,				
RA	Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,				
RA	Jones S.J., Marra M.A.;				
RT	"Generation and initial analysis of more than 15,000 full-length human				
RT	and mouse cDNA sequences."				
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Kidney;				
RA	Strausberg R.;				

RL	Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR	EMBL; BC000660; AA00660.1; -
DR	GO; GO:0005634; C:nucleus; IEA.
DR	GO; GO:0003700; P:transcription factor activity; IEA.
DR	GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR	InterPro; IPR001346; IRF.
DR	Pfam; PF00605; IRF; 1.
DR	ProDom; PD002355; IRF; 1.
DR	SMART; SM00348; IRF; 1.
DR	PROSITE; PS00601; IRF; 1.
SQ	SEQUENCE 452 AA; 49134 MW; 96B059A029751B64 CRC64;
Alignment Scores:	
Pred. No.:	1.95e-132 Length: 452
Score:	2458.50 Matches: 445
Percent Similarity:	98.45% Conservative: 0
Best Local Similarity:	98.45% Mismatches: 2
Query Match:	91.09% Indels: 6
DB:	2 Gaps: 1
US-09-975-253-1 (1-1426) x Q7Z5G6 (1-452)	
QY	47 ATGGGAACCCCAAGCAGCCAGGTCCTCCCTGGCTGGTGTGCGAGCTGGACCTGGGCGCAA 106
Db	1 MetGlyThrProLysProArgIleLeuProThrLeuValSerGlnLeuAspLeuGln 20
QY	107 CTGGAGGGGTGGCTGGGTGAACAAGAGCGCACGCGCTTCGCGATCCCTTTGGAGGAC 166
Db	21 LeuGluGlyValAlaTrpValAsnLysSerArgThrArgPheArgIleProThrLysHis 40
QY	167 GGCCTACGGCAGGATGACAGCAGGAGGATTCGGATCTTCAGAGCTGGGCGGAGGCC 226
Db	41 GlyLeuArgGlnAspAlaGlnGlnGluAspPheGlyLePheGlnAlaTrpAlaGluAla 60
QY	227 ACTGGTGCAATGTTCCCGGAGGAGTAACGACAGCTGCAACCTCGGAAGAGAAATTC 286
Db	61 ThrGlyAlaTrpValProGlyArgAspLysProAspLeuProThrLysArgAsnPhe 80
QY	287 CGCTCTCCCTCAACCCCAAGAGGGTTGCGTTTACGAGGACCGGACCAAGGACCT 346
Db	81 ArgSerAlaLeuAsnArgLysGluGlyLeuArgLeuAlaGluAspArgSerLysAspPro 100
QY	347 CACGACCCACATAAATCTAGAGTTGTGAACCTCAGGAGTGGGACTTTTCCAGCCA 406
Db	101 HisAspProHisLysIleTyrGluPheValAsnSerGlyValGlyAspPheSerGlnPro 120
QY	407 GACACCTCTCGGACACCAATGTGGAGGAGTACTTCTGATACCCAGGAGACATCTG 466
Db	121 AspThrSerProAspThrAsnGlyGlySerThrSerAspThrGlnGluAspIleLeu 140
QY	467 GATGAGTTACTGGGTAAACATGTTGGTGGCCCACTCCAGATCCGGGACCCCAAGCTG 526
Db	141 AspGluLeuLeuGlyAsnMetValLeuAlaProLeuProAspProGlyProProSerLeu 160
QY	527 GCTGTAGCCCTGAGCCCTGCCCTCAGCCCTCGGAGCCCACTGGCAATCCCACT 586
Db	161 AlaValAlaProGluProCysProGlnProLeuArgSerProSerLeuAspAsnProThr 180
QY	587 CCCTTCCCAACCTGGGCGCTCTGAGAACCCACTGAAGCGGCTGTGGTCCCGGGGAA 646
Db	181 ProPheProAsnLeuGlyProSerGluAsnProLeuLysArgLeuLeuValProGlyGlu 200
QY	647 GAGTGGGAGTTGAGGTGACAGCCCTTACCGGGGCGCCCAAGTCTCCAGACACCATC 706
Db	201 GluTrpGluPheGluValThrAlaPheTyrArgGlyArgGlnValPheGlnThrIle 220
QY	707 TCCTCCCGGAGGCTCGGCTGGTGGGTCCGAGTGGGAGGAGCAGCAGCTGCTGGA 766
Db	221 SerCysProGluGlyLeuArgLeuValGlySerGluValGlyAspArgThrLeuProGly 240
QY	767 TGGCCAGTCACACTGCGACACCTGGCATGTCCCTGACAGACAGGAGGAGTGATGACTAC 826
Db	241 TrpProValThrLeuProAspProGlyMetSerLeuThrAspArgGlyValMetSerTyr 260
QY	827 GTGAGGCATGTGCTGAGCTGCTGGTGGGGGAACTGGTCTCTGCGGGCGCGGCGAGTGG 886
Db	261 ValArgHisValLeuSerCysLeuGlyGlyGlyLeuAlaLeuTrpArgAlaGlyGlnTrp 280
QY	887 CTCTGGGCCCACGCGCTGGGCACTGCGACATACATCTGGGAGTGGAGCAGTGTCTC 946
Db	281 LeuTrpAlaGlnArgLeuGlyHisCysHisThrTyrTrpAlaValSerGlyGluLeuLeu 300
QY	947 CCCAACAGCGGGCATGGCCCTGATGGCAGGTCCCAAGGTCACCAAGGCAAGGAGGCTGTT 1006
Db	301 ProAsnSerGlyHisGlyProAspGlyGluValProLysAspLysGluGlyGlyValPhe 320
QY	1007 GACCTGGGCGCTTCAATTGT -----AGATCTGATTACTTCAACGGAAGG 1050
Db	321 AspLeuGlyProPheIleValGlySerTrpAlaProArgSerAspTyrLysHisGlyArg 340
QY	1051 AAGCGGACGCTCACACGCTATGCTCTGGTCTGTGGGGGAGTCAATGCGCCAGGA 1110
Db	341 LysArgThrLeuThrThrLeuCysProLeuValLeuCysGlyGlyValMetAlaProGly 360
QY	1111 CCAGCGGTGGACCAAGAGCTCGTGTCAAGTTGTGCCACAGTGCCTCAGGCGCTT 1170
Db	361 ProAlaValAspGlnGluAlaArgAspGlyGlnGlyCysAlaHisValProGlnGlyLeu 380
QY	1171 GTTAGAATGGCCCGGTAGGGGTGCTCTCTCCCTCGGAGAATACTGTGGA-CTGCACAT 1230
Db	381 GlyArgAsnGlyProGlyArgGlyCysLeuLeuProGlyGluTyrCysGlyPr-alaHis 400
QY	1231 TTCCAACAGCCACCCACTCTCCCTCACCTCCGACAGTACAGGCTTACTCAGGACTT 1290
Db	401 PheGlnGlnProProThrLeuProHisLeuArgProValGlnGlyLeuProAlaGlyLeu 420
QY	1291 GGTGGAGGCATGATTTCCAGGGCCCTGGGAGAGCTGAGCCCTCGCTCTCATGTGT 1350
Db	421 GlyGlyGlyHisGlyPheProGlyProThrPheGlyGluLeuSerProArgSerTrpCys 440
QY	1351 GCCTCCAAACCCCTGTCTCCACACCTCAACCAA 1386
Db	441 AlaSerAsnProProValProHisLeuAsnGln 452
RESULT 2	
Q96GL3	PRELIMINARY; PRT; 452 AA.
ID	Q96GL3
AC	Q96GL3; 2001 (TrEMBLrel. 19, Created)
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE	IRF3 protein.
OS	Homo sapiens (Human)
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCBI_TaxID=9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Brain;
RA	MEDLINE=22388257; PubMed=12477932;
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haiech F.,
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,
RA	Brownstein M.J., Uesdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA	Rana S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA	Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA	Krzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,

Jones S.J., Marza M.A.;
 "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences";
 Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 [2]

SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Strausberg R.;
 RL EMBL; BC009395; AA009395.1; -;
 DR HSP; P23906; 21RF.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001346; IRF.
 DR InterPro; IPR008984; SMAD_FHA.
 DR InterPro; IPR009058; Wing_hlx_DNA_bnd.
 DR Pfam; PF00605; IRF; 1.
 DR PRINTS; PD0267; INTFRNREGFCT.
 DR PRODOM; PD002355; IRF; 1.
 DR SMART; SM00348; IRF; 1.
 DR PROSITE; PS00601; IRF; 1.
 SQ SEQUENCE 452 AA; 49120 MW; 96B059A028AE6B64 CRC64;

Alignment Scores:
 Pred. No.: 2,896-132 Length: 452
 Score: 2455.50 Matches: 444
 Percent Similarity: 98.45% Conservative: 1
 Best Local Similarity: 98.23% Mismatches: 2
 Query Match: 90.98% Indels: 6
 DB: 2 Gaps: 1

US-09-975-253-1 (1-1426) x Q96GL3 (1-452)

QY 47 ATGGGAACCCCAAGCCAGGNTCTCTGCTGGTGTGCGAGCTGGACCTGGGGCAA 106
 Db 1 MetGlyThrProLysProArgLeuLeuProTrpLeuValSerGlnLeuAspLeuGlyGln 20
 QY 107 CTGGAGGCGTGGCCCTGGTGAACAGAGCCGCGCTTCGGATCCCTTGGGAAGCAC 166
 Db 21 LeuGluGlyValAlaTrpValAsnLysSerArgThrArgPheArgIleProTrpLysHis 40
 QY 167 GGCCTACGCGAGGATGACAGCAGGAGGATTTCGGAACTCTCCAGCCCTGGCCGAGGCC 226
 Db 41 GlyLeuArgGlnAspAlaGlnGlnGluAspPheGlyIlePheGlnAlaTrpAlaGluAla 60
 QY 227 ACTGGTGCTATGTTCCCGGGAGGATAGCCAGACCTGCCAACCTGGAGAGGAATTC 286
 Db 61 ThrGlyAlaTyrValProGlyArgAspLysProAspLeuProThrTrpLysArgAsnPhe 80
 QY 287 CGCTCTGCCCTCAACCGCAAGAGGTTGCGTTAGCAGAGACCGGAGCAAGACCT 346
 Db 81 ArgSerAlaLeuAsnArgLysGluGlyLeuArgLeuAlaGluAspArgSerLysAspPro 100
 QY 347 CACGACCCACATAAATCTACAGTTGTGAAGTTCAGGAGTTGGGACTTTTCCAGGCA 406
 Db 101 HisAspProHisLysIleTyrGluPheValAsnSerGlyValGlyAspPheSerGlnPro 120
 QY 407 GACACTCTCCGGACACCAATGGTGGAGCAGTACTCTGATACCCAGGAGACATCTTG 466
 Db 121 AspThrSerProAspThrAsnGlyGlySerThrSerAspThrGlnGluAspIleLeu 140
 QY 467 GATGAGTTACTGGGTAACTAGTGTGGCCCACTCCAGATCCGGGACCCCAAGCCTG 526
 Db 141 AspGluLeuLeuGlyAsnMetValLeuAlaProLeuProAspProGlyProProSerLeu 160
 QY 527 GCTGTAGCCCTTGAGCCCTTGCCCTCAGCGCCCTGGCGAGCCCGAGCTTGACAAATCCCACT 586
 Db 161 AlaValAlaProGluProCysProGlnProLeuArgSerProSerLeuAspAsnProThr 180
 QY 597 CCCTTCCCAACCTGGGGCCCTCTGAGAACCCACTGAGAGCGGCTGTGGTGGCGGGGAA 646
 Db 181 ProPheProAsnLeuGlyProSerGluAsnProLeuLysArgLeuLeuValProGlyGlu 200

QY 647 GAGTGGAGTTCGAGGTGACAGCCTTCTACCGGGCCGCCAAGTCTTCCAGACACCATC 706
 Db 201 GluTrpGluPheGluValThrAlaPheTyrArgGlyArgGlnValPheGlnGlnThrIle 220
 QY 707 TCCTGCCCGGAGGCGCTGGCGTGGTGGTCCGAAGTGGGAGACAGAGCGTGCCTGGA 766
 Db 221 SerCysProGluGlyLeuArgLeuValGlySerGluValGlyAspArgThrLeuProGly 240
 QY 767 TGGCCAGTCACTGCGCAGACCTGCGATGCTCCCTGACACACAGGAGGAGTATGAGTAC 826
 Db 241 TrpProValThrLeuProAspProGlyMetSerLeuThrAspArgGlyValMetSerTyr 260
 QY 827 GTGAGCATGTGCTGAGCTGCTGGTGGGGGACTGGCTCTCTGGGGGGCCGGCAGTGG 886
 Db 261 ValArgHisValLeuSerCysLeuGlyGlyGlyLeuAlaLeuTrpArgAlaGlyGlnTrp 280
 QY 887 CTCTGGGGCCAGCGGTGGGGCACTGCCACATATCTGGGCACTGAGCGAGGAGTGTCTC 946
 Db 281 LeuTrpAlaGlnArgLeuGlyHisCysHisThrTyrTrpAlaValSerGluGluLeuLeu 300
 QY 947 CCCAAGCAGCGGCGATGCGGCTGATGCGGAGTCCCAAGGACCAAGGAAGGAGCGCTGTT 1006
 Db 301 ProAsnSerGlyHisGlyProAspGlyGluValProLysAspLysGluGlyValPhe 320
 QY 1007 GACCTGGGCGCTTCAATTGT-----AGATCTGATTACCTTACCGGAAGG 1050
 Db 321 AspLeuGlyProPheIleValGlySerTrpAlaProArgSerAspTyrLeuHisGlyArg 340
 QY 1051 AAGCGACGCTCACACGCTATGCCCTCTGGTTCTGTGGGGAGTCAATGGCCCGCAGGA 1110
 Db 341 LysArgThrLeuThrThrLeuCysProLeuValLeuCysGlyGlyValMetAlaProGly 360
 QY 1111 CCAGCGCTGGACCAAGAGGCTGCTGATGCTCAAGGTTGTGCCACAGCTGCTCAGGCGCTT 1170
 Db 361 ProAlaValAspGlnGlnAlaArgAspGlyGlnGlyCysAlaHisValProGlnGlyLeu 380
 QY 1171 GGTAGAAATGGCCCGGTAGGGGTGCTCTCTCCCTGGAGAATACTGTGGACCTGCACAT 1230
 Db 381 GlyArgAsnGlyProGlyArgGlyCysLeuLeuProGlyGlyTyrCysGlyProAlaHis 400
 QY 1231 TTCCACACGACCCACTCTCCCTCAGTCCGACAGTACAAAGGCTACTCGAGGACTT 1290
 Db 401 PheGlnGlnProProThrLeuProHisLeuArgProValGlnGlyLeuProAlaGlyLeu 420
 QY 1291 GGTGAGGCGATGGATTTCAGGGCCCTGGGAGAGCTGAGCCCTCGCTCCTCCTCGTGT 1350
 Db 421 GlyGlyGlyHisGlyPheProGlyProTrpGlyAspLeuSerProArgSerTrpCys 440
 QY 1351 GCCTCCAAACCCCTGTTTCCCAACCCACCTCAACCAA 1386
 Db 441 AlaSerAsnProProValProHisLeuAsnGln 452

RESULT 3

IRF3_HUMAN

ID IRF3_HUMAN STANDARD; PRT; 427 AA.
 AC Q14653;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-OCT-2004 (Rel. 45, Last annotation update)
 DE Interferon regulatory factor 3 (IRF-3).

GN Name=IRF3;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Retina;

RA MEDLINE=96102173; PubMed=8524823;

RX Au W.-C., Moore P.P.A., Lowther W.W., Juang Y.-T., Pitha P.M.;

RT "Identification of a member of the interferon regulatory factor family

that binds to the interferon-stimulated response element and activates

RT expression of interferon-induced genes.";
 RL Proc. Natl. Acad. Sci. U.S.A. 92:11657-11661(1995).
 RN [2]
 RP SEQUENCE OF 323-413 FROM N.A.
 RX MEDLINE=99020108; PubMed=9803267;
 RA Bellingham J., Gregory-Evans K., Gregory-Evans C.Y.; to chromosome
 RT "Mapping of human interferon regulatory factor 3 (IRF3) to chromosome
 RL 19q13.3-13.4 by an intragenic polymorphic marker.";
 RL Ann. Hum. Genet. 62:231-234(1998)
 CC Ann. Hum. Genet. 62:231-234(1998)
 CC other transcription factors, possibly members of the SPAT
 CC family. Binds specifically to the IFN-stimulated response element
 CC (ISRE) but not to the IRF-1 binding site PRD-1.
 CC -! SUBCELLULAR LOCATION: Nuclear.
 CC -! TISSUE SPECIFICITY: Expressed constitutively in a variety of
 CC tissues.
 CC -! SIMILARITY: Belongs to the IRF family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z56281; CAA91227.1; -;
 DR EMBL; U86636; AAC68818.1; -;
 DR PDB; 1J2F; X-ray; A/B=170-427.
 DR PDB; 1QWT; X-ray; A/B=173-427.
 DR TRANSFAC; T04673; -;
 DR Genew; HGNC:6118; IRF3.
 DR MIM; 603734; -;
 DR GO; GO:0003702; F:RNA polymerase II transcription factor acti. . .; TAS.
 DR GO; GO:0003712; P:transcription cofactor activity; TAS.
 DR GO; GO:0006366; P:transcription from Pol II promoter; TAS.
 DR InterPro; IPR001346; IRF.
 DR InterPro; IPR008984; SMAD FHA.
 DR InterPro; IPR009058; Wing_hlx_DNA_bnd.
 DR Pfam; PF00605; IRF_1
 DR PRINTS; PR02667; INTFRNREGFCT.
 DR ProDom; PD002355; IRF; 1.
 DR SMART; SM00348; IRF; 1.
 DR PROSITE; PS00601; IRF; 1.
 DR 3D-structure; Activator; DNA-binding; Nuclear protein; Polymorphism;
 KW Transcription regulation.
 FT DNA BIND 7 107 Tryptophan pentad repeat.
 FT VARIANT 96 96 R -> Q (in dbSNP:968457).
 FT VARIANT 377 377 E -> K (in dbSNP:1049486).
 FT VARIANT 427 427 S -> T (in dbSNP:7251).
 FT VARIANT 427 427 /FTId=VAR_011902.
 FT VARIANT 427 427 /FTId=VAR_011903.
 SQ SEQUENCE 427 AA; 47219 MW; F536676FA78B0110 CRC64;
 Alignment Scores:
 Pred. No.: 7,21e-125 Length: 427
 Score: 2326.00 Matches: 426
 Percent Similarity: 99.77% Conservative: 0
 Best local Similarity: 99.77% Mismatches: 1
 Query Match: 86.18% Indels: 0
 DB: 1 Gaps: 0
 US-09-975-253-1 (1-1426) x IRF3_HUMAN (1-427)
 QY 47 ATGGGAACCCCAAGCCAGGCTCTCGCTGGTGTTCGAGCTGAGCTGGGCGAA 106
 Db 1 MetGlyThrProLysProArgLeuLeuProThrLeuValSerGlnLeuAspLeuGln 20
 QY 107 CTGGAGGGCGTGGCTGGTGAACAGAGCCGAGCCGCTTCGCGATCCCTTGGAGGAC 166
 Db 21 LeuGluGlyValAlaTrpValAsnLysSerArgThrArgPheArgLeuProThrLysHis 40

QY 167 GGCCTACGGCAGGATGCACAGCAGGAGGATTTCCGAAATCTTCCAGGCTGGCGGAGCC 226
 Db 41 GlyLeuArgGlnAspAlaGlnGlnGluAspPheGlyLeuPheGlnAlaTrpAlaGluAla 60
 QY 227 ACTGGTCATATGTTCCCGGAGGAGGATTAAGCCAGACCTGCCCAACTGGAAGAGAAATTC 286
 Db 61 ThrGlyAlaTrpValProGlyArgAspLysProAspLeuProThrTrpLysArgAsnPhe 80
 QY 287 CGCTCTGCCCTCAACCGCAAGAGGGTGGTGTTCAGAGAGGACCGGAGCAAGACCCCT 346
 Db 81 ArgSerAlaLeuAsnArgLysGluGlyLeuArgLeuAlaGluAspArgSerLysAspPro 100
 QY 347 CACGACCCACATAAATACAGATTGTGAACCTCAGAGTTGGGGACTTTTCCAGGCA 406
 Db 101 HisAspProHisLysIleTyrGluPheValAsnSerGlyValGlyAspPheSerGlnPro 120
 QY 407 GACACTCTCCGACACCAATGGTGGAGGACGACTTCTGTATACCCAGGAGACATTCG 466
 Db 121 AspThrSerProAspThrAsnGlyGlyGlySerThrSerAspThrGlnGluAspLeu 140
 QY 467 GATGATTACTGGGTAAACATGTGTGGCCCACTCCACAGATCCCGGAGCCCGAGCCCTG 526
 Db 141 AspGluLeuLeuGlyAsnMetValLeuAlaProLeuProAspProGlyProSerLeu 160
 QY 527 GCTGTAGCCCTGAGCCCTGCGCTCAGCCCTGCGGAGCCCGAGCTTGGACATCCCACT 586
 Db 161 AlaValAlaProGluProCysProGlnProLeuArgSerProSerLeuAspAsnProThr 180
 QY 587 CCTTCCCAAACTGGGGCCCTCTGAGAACCCCACTGAAGCGGCTGTGTGGCGGGGAA 646
 Db 181 ProPheProAsnLeuGlyProSerGluAsnProLeuLysArgLeuLeuValProGlyGlu 200
 QY 647 GAGTGGGAGTTGAGGTGACAGCTTCTACCGGGCGCCCAAGTCTTCCAGCAGACCATC 706
 Db 201 GluTrpGluPheGluValThrAlaPheTyrArgGlyArgGlnValPheGlnGlnThrIle 220
 QY 707 TCCTGCGCGGAGGCGCTGGGCTGGTGGGTTCGAAATGGGAGACAGGAGCGCTGGGA 766
 Db 221 SerCysProGluGlyLeuArgLeuValGlySerGluValGlyAspArgThrLeuProGly 240
 QY 767 TGGCCAGTCACCTGCGCAGACCCCTGCGCATCTCCTGACAGACAGAGGAGGTAGAGTAC 826
 Db 241 TrpProValThrLeuProAspProGlyMetSerLeuThrAspArgGlyValMetSerTyr 260
 QY 827 GTGAGGATCTGTGAGCTGCTGGTGGGAGTGTGCTCTCTGGCGGGCGGCGAGTGG 886
 Db 261 ValArgHisValLeuSerCysLeuGlyGlyLeuAlaLeuTrpArgAlaGlyGlnTrp 280
 QY 887 CTCTGGGCCCGAGGCTGGGCGACTGCCACATACCTAGTGGGAGGAGGAGGTGCTC 946
 Db 281 LeuTrpAlaGlnArgLeuGlyHisCysHisThrTyrTrpAlaValSerGluGluLeu 300
 QY 947 CCCAACAGCGGCGATGGCGTGTGGCGAGTCCCAAGGACAAAGGAGGAGGCGGTGT 1006
 Db 301 ProAsnSerGlyHisGlyProAspGlyGluValProLysAspLysGlyGlyValPhe 320
 QY 1007 GACCTGGGGCCCTTCATTGTAGATCTGATTACCTTCCAGGAAGGAGGAGCGCTCACC 1066
 Db 321 AspLeuGlyProPheLeuValAspLeuIleThrPheThrGluGlySerGlyArgSerPro 340
 QY 1067 CGCTATGCCCTCTGCTGTGGGAGTTCATGGCCCGGAGGAGGAGGAGGAGGAGGAG 1126
 Db 341 ArgTrpAlaLeuTrpPheCysValGlyGlySerTrpProGlnAspGlnProTrpThrLys 360
 QY 1127 AGGCTCGTGTGATCAAGTTGTGCCACAGCTGCTCAGGGCTTGGTGAAGATGCCCGG 1186
 Db 361 ArgLeuValMetValLysValValProThrCysLeuArgAlaLeuValGluMetAlaArg 380
 QY 1187 GTAGGGGTGCTCTCTCCCTCGAGAAATCTGTGGACCTGTCACATTTCCACAGCCACCA 1246
 Db 381 ValGlyGlyAlaSerSerLeuGluAsnThrValAspLeuHisIleSerAsnSerHisPro 400
 QY 1247 CTCTCCCTCACCCTCGACACAGTACAGAGGCTACCTGCGAGGACTTGTGTGGAGGCGATGGAT 1306

```

Db 401 LeuSerLeuThrSerAspGlnTyrLysAlaTyrLeuGlnAspLeuValGluGlyMetAsp 420
QY 1307 TTCAGGGCCCTGGGAGAC 1327
Db 421 PheGlnGlyProGlyGluSer 427

RESULT 4
ID AAH71721 PRELIMINARY; PRT; 427 AA.
AC AAH71721.
DT 01-JUN-2004 (TrEMBLrel. 27, Created)
DT 01-JUN-2004 (TrEMBLrel. 27, Last sequence update)
DE 01-JUN-2004 (TrEMBLrel. 27, Last annotation update)
DE IRF3 protein.
GN IRF3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalusz D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RA Strausberg R.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC071721; AAH71721.1; -.
SQ SEQUENCE 427 AA; 47219 MW; F536676FA78B0110 CRC64;

Alignment Scores:
Pred. No.: 7,21e-125 Length: 427
Score: 2326.00 Matches: 426
Percent Similarity: 99.77% Conservative: 0
Best Local Similarity: 99.77% Mismatches: 1
Query Match: 86.18% Indels: 0
DB: 2 Gaps: 0

US-09-975-253-1 (1-1426) x AAH71721 (1-427)

QY 47 ATGGGAACCCCAAGCCAGGNTCTCGCTCGCTGGTGTGGAGCTGGACCTGGGGCAA 106
Db 1 MetGlyThrProLysProArgGlyLeuProThrLeuValSerGlnLeuAspLeuGlyGln 20
QY 107 CTGAGGGCGTGGCTGGGTGAACAAGACGCGCTCCGCATCCCTCGAAGTGAAGCAC 166
Db 21 LeuGluGlyValAlaTpvAlaValAsnLysSerArgThrArgPheArgIleProThrLysHis 40
QY 167 GGCTACGGCAGGTGACAGCAGGAGATTTCGAATCTTCCAGGCTGGCGGAGGCC 226
Db 41 GlyLeuArgGlnAspAlaGlnGlnGluAspPheGlyIlePheGlnAlaTpvAlaGluAla 60

```

```

QY 227 ACTGTGTCATATGTTCCGGGAGGGATAAGCCAGACCTGCCAACCTGGAAGAGAAATTC 286
Db 61 ThrGlyAlaTyrValProGlyArgAspLysProAspLeuProThrTyrLysArgAspPhe 80
QY 287 CGCTCTGCCCTCAACCGCAAGAGGGTTGCGTTTAGCAGAGACCGGACGAGGACCT 346
Db 81 ArgSerAlaLeuAsnArgLysGluGlyLeuArgLeuAlaGluAspArgSerLysAspPro 100
QY 347 CACGACCCACATAAAATCTACGAGTTTGTGAATCTCAGGAGTTGGGACTTTTCCAGCCA 406
Db 101 HisAspProHisLysIleTyrGluPheValAsnSerGlyValGlyAspPheSerGlnPro 120
QY 407 GACACCTCTCCGGACACCAATGGTGAGGAGCTACTTCTGATACCCAGGAGACATCTG 466
Db 121 AspThrSerProAspThrAsnGlyGlySerThrSerAspThrGlnGluAspIleLeu 140
QY 467 GATGAGTTACTGGTTAACTGGTTGGTGGCCCACTCCAGATCCGGGACCCCAAGCCTG 526
Db 141 AspGluLeuLeuGlyAsnMetValLeuAlaProLeuProAspProGlyProProSerLeu 160
QY 527 GCTGTAGCCCTGAGCCCTGCCCTCAGCCCTGCGGAGCCCGAGCTTGACATCCCACT 586
Db 161 AlaValAlaProGluProCysProGlnProLeuArgSerProSerLeuAspAsnProThr 180
QY 587 CCTTCCCAACCTCGGGCCCTCTGAGAACCCACTGAGACGGCTGTGTGTCGGGGGAA 646
Db 181 ProPheProAsnLeuGlyProSerGluAsnProLeuLysArgLeuValProGlyGlu 200
QY 647 GAGTGGGAGTTCGAGGTGACAGCTTCTACGGGCGCCCAAGTCTTCAGCAGACCATC 706
Db 201 GluThrGluPheGluValThrAlaPheTyrArgGlyArgGlnValPheGlnThrIle 220
QY 707 TCTCTCCCGAGGGCTCGGGCTCGGTGGTCCGAGTGGGAGTGGAGACAGACGCTGCTGA 766
Db 221 SerCysProGluGlyLeuArgLeuValGlySerGluValGlyAspArgThrLeuProGly 240
QY 767 TGGCCAGTCACACTGCCAGACCTCGCATGTCCCTGACAGACAGGGAGTGTAGTAC 826
Db 241 TrpProValThrLeuProAspProGlyMetSerLeuThrAspArgGlyValMetSerTyr 260
QY 827 GTGAGGCATGTCTGAGTGTGGTGGGAGCTGGCTCTCTGCGCGCGCGGAGTGG 886
Db 261 ValArgHisValLeuSerCysLeuGlyGlyGlyLeuAlaLeuTrpArgAlaGlyGlnTrp 280
QY 887 CTCTGGGCGCCAGCGCTGGGCACTGCCACATCTGGGAGTGGGAGTGGAGAGTGTCTC 946
Db 281 LeuTrpAlaGlnArgLeuGlyHisCysHisThrTyrTrpAlaValSerGluLeuLeu 300
QY 947 CCCAACAGCGGCGATGGGCTGTGATGGGAGGTCCCAAGGACAAGGAGGAGGTGTT 1006
Db 301 ProAsnSerGlyHisGlyProAspGlyGluValProLysAspLysGluGlyValPhe 320
QY 1007 GACTGGGGCCCTTATTGTAGATCTGATTACCTTCAAGGAGGAGGAGCGTCAACA 1066
Db 321 AspLeuGlyProPheIleValAspLeuIleThrPheThrGluGlySerGlyArgSerPro 340
QY 1067 CGCTATGCCCTCTGTTCTGTGGGGAGTCTATGGCCCGGAGGAGGAGGAGGAGGAG 1126
Db 341 ArgTyrAlaLeuTrpPheCysValGlyGluSerTrpProGlnAspGlnProTrpThrLys 360
QY 1127 AGGCTCGTGTGTTCAAGGTTGTGCCACCTGCTCAGGGCTTGTGTAGAAATGCGCCGG 1186
Db 361 ArgLeuValMetValLysValValProThrCysLeuArgAlaLeuValGluMetAlaArg 380
QY 1187 GTAGGGGTGCTCTCTCCCTGGAGAAATCTGTGGACCTGCACATTTCCACAGCACCCA 1246
Db 381 ValGlyGlyAlaSerSerLeuGluAsnThrValAspLeuHisIleSerAsnSerHisPro 400
QY 1247 CTCTCCCTCACTCCAGCAGTACAGGCTACTCGAGGACTGTGTGGAGGCGCATGAT 1306
Db 401 LeuSerLeuThrSerAspGlnTyrLysAlaTyrLeuGlnAspLeuValGluGlyMetAsp 420

```


Fri Oct 29 11:11:45 2004

```

1307 TTCAGGCGCCCTGGGGAGAGC 1327
421 PheGlnGlyProGlyGluSer 427

RESULT 5
Q764M6 PRELIMINARY; PRT; 419 AA.
ID Q764M6
AC Q764M6; 2004 (Tremblrel. 27, Created)
DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)
DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)
DE Interferon regulatory factor 3.
GN Name=irf3;
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
NCBI_TaxID=9823;
RN [1]
SEQUENCE FROM N.A.
RA Uenishi H., Echuchi T., Suzuki K., Sawazaki T., Toki D., Shinkai H.,
RA Okumura N., Hamashina N., Awata T.;
RT "PEDE (Pig EST Data Explorer): construction of a database for ESTs
RL Nucleic Acids Res. 32:D484-D488(2004).
DR EMBL; AB116563; BAD06317.1; -.
DR InterPro; IPR001346; IRF.
DR InterPro; IPR008984; SMAD_PHA.
DR InterPro; IPR009058; Wing_hlx_DNA_bnd.
DR Pfam; PF00605; IRF; 1.
DR PRINTS; PRO0267; INTERFERGFC.
DR ProDom; PD002355; IRF; 1.
DR SMART; SM00348; IRF; 1.
DR PROSITE; PS00601; IRF; UNKNOWN 1.
SQ SEQUENCE 419 AA; 46645 MW; 8B731861BE002284 CRC64;

Alignment Scores:
Pred. No.: 1.68e-91 Length: 419
Score: 1742.00 Matches: 327
Percent Similarity: 83.61% Conservative: 25
Best Local Similarity: 77.67% Mismatches: 67
Query Match: 64.54% Indels: 2
DB: 2 Gaps: 1

US-09-975-253-1 (1-1426) x Q764M6 (1-419)
QY 47 ATGGGAACCCCAAGCCAGCGTCTCGCTGGCTGGTGTGGAGCTGGAGCTGGGCGAA 106
Db 1 MetGlyThrGlnLysProArgGileLeuProTrpLeuLeuSerGlnLeuAsnGlnGlyGln 20
QY 107 CTGAGGCGCGTGGCTGGTGAACAAGAGCGCGACCGCTTCCGCATCCCTTGAAGCAC 166
Db 21 LeuGluGlyValAlaTrpLeuAspGluGlyHisThrArgPheArgGileProTrpLysHis 40
QY 167 GGCCTACGGCAGGATGACAGCAGGAGATTTCGGATCTTCCAGGCTGGCGCGAGGCC 226
Db 41 GlyLeuArgGlnAspAlaGlnGlnGluAspPheGlyPheGlnAlaTrpAlaGluAla 60
QY 227 ACTGGTGCATATGTTCCCGGAGGAGTAAGCCAGACCTGCCAACCTGGAAGAGGAATTC 286
Db 61 SerGlyAlaIleThrProGlyLysAspLysProAspLeuProThrTrpLysArgAsnPhe 80
QY 287 CGCTCGCCCTCAACCGCAAGAGGGTTGGCTTTAGCAGAGGACCGGAGGACCGCT 346
Db 81 ArgSerAlaLeuAsnArgLysGluAlaLeuArgLeuAlaGluAspHisSerLysAspPro 100
QY 347 CACGACCCACATAAATCTACGAGTTGTGAAGTGGAGTTGGGACTTTTCCAGGCCA 406
Db 101 HisAspProHisLysIleThrGluPheValThrSerGlyValGlyAspPheProGluPro 120
QY 407 GACACTCTCCGACACCACTGGTGGAGGACAGTACTTCTGTATACCCAGAGCAATCTTG 466
Db 121 AspThrSerLeuAspLeuSerGlyArgTyrSerThrSerAspThrHisGluAspSerLeu 140

467 GATCAGTTACTGGTAACTAGTGTGGTGGCCCTCCAGATCCGGGAGCCCAAGCGCTG 526
141 AsplysLeuLeuSerGlyMetAspLeuA-----SerAspAlaGlyProGlnSerLeu 158
527 GCTGTAGCCCTGAGCCCTGCCCTCAGCCCTCGGGAGCCCAAGTTCGACAAATCCCACT 586
159 ThrLeuAlaLeuGlnProGlnLeuSerLeuSerProSerValAspAlaProAla 178
587 CCCTTCCCAACCTGGGCGCCCTCTGAGAACCCTGAAGCGGCTGTGGTCCCGGGGAA 646
179 SerCysProLeuLeuGlyValArgGluAsnProLeuLysGlnLeuLeuAlaAsnAsp 198
647 GAGTGGGAGTTCGAGGTGACAGCCCTTACCGGGGCGCCCAAGTTCCTCCACAGACCATC 706
199 GlnTrpGluPheGlnValThrValPheTyrArgGlyCysGlnValPheGlnGlnThrVal 218
707 TCCTGCCCGGAGGCGCTGGCTGGTGGTCCGAAGTGGGAGAGCAGACAGCTGCTGGA 766
219 CysSerProGlyGlyLeuArgLeuValGlySerGluAlaGluAspGlyThrLeuAlaGly 238
767 TGGCCAGTCACACTGCCAGACCCCTGGCATGTCCTGTGACAGAGGGGAGTGTAGCTAC 826
239 GlnProValArgLeuProAspProAlaAlaSerLeuThrAspArgGlyValAlaAspTyr 258
827 GTGAGGCGATGCTGAGCTGCTGGTGGGAGTGGCTCTCTGCGGGCGCGGAGTGG 886
259 ValArgArgValLeuSerCysLeuGlyGlyLeuAlaLeuTrpArgAlaGlyGlnTrp 278
887 CTCTGGGCGCGGCTGGGCGCTGCCACATCTGGGCGAGTGGGAGGAGGAGTGTCTC 946
279 LeuTrpAlaGlnArgLeuGlyHisCysHisValTyrTrpAlaMetGlyGluGluLeu 298
947 CCCACAGCGCGCTGGGCTGATGCGGAGTGGTCCCAAGCAAGCAAGAGGAGCGCTGTT 1006
299 ProAspSerGlyHisLysProAspGlyGluValProLysAspArgGluGlyValPhe 318
1007 GACCTGGGCGCTTCTATTGTAGTCTGATCTTCCAGGAGGAGGAGCGCTCACCA 1066
319 AspLeuGlyProPheIleGluAspLeuAlaPheIleGluGlySerArgArgSerPro 338
1067 CGCTATGCTGCTGGTCTGTGGGAGTGTATGCGCCAGGAGCGCGCTGGAGCAAG 1126
339 ArgTyrThrLeuTrpPheCysMetGlyGlnSerTrpProGlnAspGluProTrpValLys 358
1127 AGGCTCGTATGCTGAAGTGTGTGCCACCTGCTGAGGCGCTGTGTAGAAATGGCCCG 1186
359 ArgLeuValMetValLysValValProMetCysLeuArgAlaLeuValAspMetAlaArg 378
1187 GTAGGGGCTGCTCTCTCCCTGGAGATATCTGTGACCTGCACATTTTCCACAGCCCA 1246
379 AspGlyGlyAlaSerSerLeuGluAsnThrValAspLeuHisLysSerAsnSerHisPro 398
1247 CTCTCCCTCACCTCCGACCACTGACAGGCTTACCTGACGAGCTTGTGGAGGGATGGAT 1306
399 LeuSerLeuThrSerAspGlnTyrLysAlaCysLeuArgAspLeuValGluAspMetAsp 418
1307 TTC 1309
419 Phe 419

RESULT 6
BAD06317 PRELIMINARY; PRT; 419 AA.
ID BAD06317;
AC BAD06317;
DT 02-MAR-2004 (Tremblrel. 27, Created)
DT 02-MAR-2004 (Tremblrel. 27, Last sequence update)
DE Interferon regulatory factor 3.
GN IRF3.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
NCBI_TaxID=9823;

```


Schmerch A., Schein J.E., Jones S.J.M., Marra M.A.;
"generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
-!- FUNCTION: May activate transcription by complex formation with
other transcriptional factors, possibly members of the STAT
family. Binds specifically to the IFN-stimulated response element
(ISRE) but not to the IRF-1 binding site PRD-1 (By similarity).
CC (ISRE) but not to the IRF-1 binding site PRD-1 (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: Belongs to the IRF family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; U75839; AAB36924.1; --
CC EMBL; U75840; AAB36925.1; --
CC EMBL; BC050882; AAH50882.1; --
CC HSPG; P23906; 2IRF.
CC MGD; MGI:1859179; Irf3.
CC InterPro; IPR001346; IRF.
CC InterPro; IPR009884; SWAD FHA.
CC InterPro; IPR009058; Wngd_hlx_DNA_bnd.
CC Pfam; PF00605; IRF; 1.
CC PRINTS; PR00267; INTFRNREGFCT.
CC ProDom; PD002355; IRF; 1.
CC PROSITE; PS00601; IRF; 1.
CC Activator; DNA-binding; Nuclear protein; Transcription regulation.
CC DNA BIND 7 107
CC ORFANCE 419 A.B. 46952 MW: 1PF67C4E0FC7F027 CRC64;
CC ORFANCE 419 A.B. 46952 MW: 1PF67C4E0FC7F027 CRC64;

Alignment Scores:	Length:	419
Pred. No.:	1,064-80	301
Score:	1553.00	33
Percent Similarity:	78.59%	83
Best Local Similarity:	70.82%	8
Query Match:	57.54%	6
DB:	1	

US-09-975-253-1 (1-1426) x IRF3_MOUSE (1-419)	
QY 47 ATGGGAACCCCAAAGCCACGGNTCTCGCCCTGGCTGGTGTCCAGCTGCACCTGGGGCAA	106
Db 1 MetGluThrProIysProArgIleLeuProTPrLeuValSerGlnLeuAspLeuGlyGln	20
QY 107 CTGGAGGGCGTGGCCCTGGTGAAACAGACCGCAGCGGCTTCGGATCCCTGGAGGACAC	166
Db 21 LeuGluGlyValAlaTrpLeuAspGluSerArgThrArgPheArgIleProTrpIysHis	40
QY 167 GGCCTACGGCAGGATGCACAGCAGGAGGATTTCCGAATCTTCAGGCCTGGCGCAGGCGC	226
Db 41 GlyLeuArgGlnAspAlaGlnMetAlaAspPheGlyIlePheGlnIleTrpAlaGluAla	60
QY 227 ACTGGTGCATATGTTCCCGGGAGGGATAGCCAGACCTGCACACCTGGAGAGGAATTC	286
Db 61 SerGlyAlaIleThrProGlyIysAspLysProAspValSerThrTrpIysArgAsnPhe	80
QY 287 CGCTCTGCCCTCAACGCCAAAGAGGGTTGCGTTTAGCAGAGACCGGAGCAGGACCCCT	346
Db 81 ArgSerAlaLeuAsnArgLysGluValLeuArgLeuAlaAlaAspAsnSerIysAspPro	100
QY 347 CACGACCCCAATAAAATCTACGAGTTTGTAATCTCAGGAGTTGGGACTTTTCCAGCCA	406
Db 101 TyrAspProHisLysValTyrGluPheValThrProGlyAlaArgAspPheValHisLeu	120
QY 407 GACACCTCTCCGACACCAATGGTGGAGCGAGTACTTCTCATACCCAGGAGACATTCTG	466
Db 121 GlyAlaSerProAspThrAsnGlyLysSerSerLeuProHisSerGlnGluAsn--Leu	139

1120 GACCAAGAGCTCGTATGCTCAAGTTGTGCCACGCTGCTCAGGCGCTTGTAGAAAT 1179
 404 pAsnLysIlePheIleMetValGlnIleValProValSerMetLysLeuLeuHisAspMe 424
 1180 GGCCCGGGTGGGGTGGCTCTCCCTGGAGAACTACT---GTGGACCTGCACATTTCCAA 1236
 424 tSerTyrSerThrGlyAlaSerSerLeuHisSerGluIleAsnLeuGluIleSerAs 444
 1237 CAGCACCCACTCTCCCTCACCCTCCGACCACTCAAGGCCTACCTGCAGGACTTTGGTGA 1296
 444 pSer-----LeuSerSerThrAsnAsp---ValMetAlaValLeuArgGluLeuHisG1 461
 1297 GGGCATGGATTCCAG 1312
 461 uMetMetAspPheGlu 466

RESULT 9
 Q6RCM8
 ID Q6RCM8 PRELIMINARY; PRT; 488 AA.
 AC Q6RCM8; (TREMELrel. 27, Created)
 DT 05-JUL-2004 (TREMELrel. 27, Last sequence update)
 DT 05-JUL-2004 (TREMELrel. 27, Last annotation update)
 DE Interferon regulatory factor 5 variant 4 (Interferon regulatory factor
 5 variant 3).
 GN Name=IRF5;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP MEDLINE FROM N.A.
 RX MEDLINE=21316468; PubMed=11303025;
 RA Barnes B.J., Moore P.A., Pitha P.M.;
 RT "virus-specific activation of a novel interferon regulatory factor,
 RT IRF-5, results in the induction of distinct interferon alpha genes.";
 RL J. Biol. Chem. 276:23382-23390(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Barnes B.J., Pitha P.M.;
 RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY504947; AAR90326.1; -
 DR EMBL; AY504946; AAR90325.1; -
 DR InterPro; IPR001346; IRF.
 DR InterPro; IPR008984; SMAD FHA.
 DR InterPro; IPR009058; Wing_hlx_DNA_bnd.
 DR Pfam; PF00605; IRF; 1.
 DR PRINTS; PR00267; INTERFERGCT.
 DR ProDom; PD002355; IRF; 1.
 DR SMART; SM00348; IRF; 1.
 DR PROSITE; PS00601; IRF; 1.
 DR SEQUENCE 488 AA; 54942 MW; 4B43B54BE74A7E37 CRC64;

Alignment Scores:
 Pred. No.: 8.74e-21 Length: 488
 Score: 504.50 Matches: 163
 Percent Similarity: 44.47% Conservative: 62
 Best Local Similarity: 32.21% Mismatches: 186
 Query Match: 18.69% Indels: 95
 DB: 2 Gaps:

US-09-975-253-1 (1-1426) x Q6RCM8 (1-488)
 QY 2 GTTCCAGCTGCCGCGACCCGCGACCTTCCATCGTAGGCGGACCATGGGAACCCCAAG 61
 Db 5 IleProValAlaProThrProArgVal----- 15
 QY 62 CCACGGNCTCCTCCCTGGTGTGTCGACGTGGAGCTGGGGCACTGGAGGGCGTGGCC 121
 Db 16 ---ArgLeuPyrProThrLeuValAlaGlnValAsnSerCysGlnTyrProGlyLeuGln 34
 QY 122 TGGGTAAACAGACGCGCGCTTCCGCTCCCTCCGCTTGGGAAGCAGCGCGCTACGGCAGGAT 181

35 TrpValAsnGlyGluLysLysLeuPheCysIleProTrpArgHisAlaThrArgHisGly 54
 182 GCACAGCAGGAG---GATTTTCGGAATCTTCAGCGCTGGCGCCAGGCCACTGTGTGCATAT 238
 55 ProSerGlnAspGlyAspAsnThrIlePheLysAlaTrpAlaLysGluThrGlyLysTyr 74
 239 GTTCCCGGGGAGGATAAGCCACCTGCCAACCTGGAGAGGAAATTTCCGCTCTGCCCTC 298
 75 ThrGluGlyValAspGluAlaAspProAlaLysTrpLysAlaAsnLeuArgCysAlaLeu 94
 299 AACCGCAAGAGAGGTTCGTTTACGAGACCGGAGCAAGGAC---CCTCAGCACCA 355
 95 AsnLysSerArgAspPheArgLeuIleTyrAspGlyProArgAspMetProProGlnPro 114
 356 CATAAATCTACGAGTTTGTGAACCTCAGGAGTGTGGGACTTTTCCAGCCAGCACCTCT 415
 115 TyrLysIleTyrGluValCysSerAsnGlyPro-----AlaProThrAspSerGln 131
 416 CCG-----GACACCAATGTTGGAGGAGTACTTCTGATACCCAGCAAGACATTCGGAT 469
 132 ProProGluAspTyrSerPheGlyAlaGlyGluGluGluGluGluGluGluGln 151
 470 GAGTTACTGGGTAAACATG-----GTGTTGCCCA 499
 152 ArgMetLeuProSerLeuSerLeuThrGluAspValLysTrpProProThrLeuGlnPro 171
 500 CTCACGATCCCGGAGCCCGGAGCTGTAGCCCTGAGCCCTGAGCCCTCAGCCCTG 559
 172 ---ProThrLeuGlnProProValLeuGlyProProAlaProAspSerProLeu 190
 560 CGGAGCCCGAGCTTGACAAATCCACTCCCTCCCAACCTGGGCGCTCTGAG----- 613
 191 ---AlaProProGlyAsnProAlaGlyPheArgGluLeu---LeuSerGluValLeu 208
 614 -----AACCCACTGAAGCGCTTGTGTGCGCC---GGGCAAGAG----- 649
 209 GluProGlyProLeuProAlaSerLeuProProAlaGlyGluGlnLeuLeuProAspLeu 228
 650 -----TGGGATTCGAGTGACAGCTTCTACCGGGC 682
 229 LeuLeuSerProHisMetLeuProLeuThrAspLeuGluLeuLysPheGlnTyrArgGly 248
 683 CGCCAAAGTCTTCCAGCAGACCATCTCTCCCGGAGGCGCTCGGCTGTGGTCCGAA 742
 249 ArgProProArgAlaLeuThrIleSerAsnProHisGlyCysArgLeuPheTyrSerGln 268
 743 GTG-----GGAGACAGGAGCGCTGCTCGTGGATGCGCA 772
 269 LeuGluAlaThrGlnGluGlnValGluLeuPheGlyProIleSerLeuGlu-----Gln 286
 773 GTCACACTGCCAGACCCCTGGCATGTCCTCAGACAGAGGGGAGTGATGACTAGTGAGG 832
 287 ValArgPheProSerProGluAspIleProSerAspLys---GlnArgPheTyrThrAsn 305
 833 CATGTGTGAGCTGCTGGTGGGGGAGTGTCTCTGGCGGGCGGCGAGTGTCTG 892
 306 GlnLeuLeuAspValLeuAspArgGlyLeuLeuLeuGlnGlnGlnGlnAspLeuTyr 325
 893 GCCCAGCGCTGGGCGCATCCACACATCTG-----GCATGAGCGGAGGAG 940
 326 AlaIleArgLeuCysGlnCysLysValPheTrpSerGlyProCysAlaSerAlaHisAsp 345
 941 CTGCTCCCAACAGCGGCGCATGGGCGCTGATCGGAGGTCCCGAGGACCAAGGAGGAGC 1000
 346 SerCysProAsn-----ProIleGlnArgValLysThrLys 358
 1001 GTGTTTACCTGGGCGCTTCTTATTGATGATCTGATTACCTTCCAGCAAGGA---AGCGGA 1057
 359 LeuPheSerLeuGluHisPheLeuAsnGluLeuIleLeuPheGlnLysGlyGlnThrAsn 378
 1058 CGCTCACCAGCGTATGCGCTCTGTTCTGTGGGGAGTGTATGCGCCCGAGCCAGCCG 1117
 379 ThrProProProPheGluIlePheCysPheGlyGluGluTrpProAspArgLysPro 398

QY	239	GTTC	CGGAGGGATTAAGCCAGACCTG	CCAACTGGAAAGGAAATTTCCGCTCTCCCTC	298		
DB	75	Thr	Glu	GlyValAspGluAlaAspProAla	lysTrp	lysAlaAsnLeuArgCysAlaLeu	94
QY	299	AACCG	AAAGAGGTTTCGTTTACGAGACG	CCGGACAGGAC	---CCTCAGACCA	355	
DB	95	Asn	lysSerArgAspPhe	ArgLeuLeuIleTyrAsp	GlyProArgAspMetPro	GlnPro	114
QY	356	CATAA	ATCTACGAGTTTGTGAATCAGGAGT	CGGGACTTTTCCAGCCAGACACCTCT	415		
DB	115	Tyr	lysIleTyrGluValCysSer	AsnGlyPro	-----AlaProThrAspSerGln	131	
QY	416	CCG	-----GACACCAATGTTGGAGG	CAGTACTTCTGATACCCAGGAAGACTTC	TGGAT	469	
DB	132	Pro	GluAspTyrSerPhe	GlyAlaGlyGluGluGluGluGluGluGluGln	151		
QY	470	GAGTT	ACTGGGTAACTG	-----GTGTTGGCCCCA	490		
DB	152	Arg	MetLeuProSerLeuSerLeuThr	GluAspVallylTyrProProThrLeuGlnPro	171		
QY	500	CTCC	AGATCCGGGACCCCAAGCCTGGCTGT	AGCCCTCAGCCCTCAGCCCTCAGCCCTG	559		
DB	172	---Pro	ThrLeuGlnProProValValLeuGlyProProAla	ProProAspProSerProLeu	190		
QY	560	CGGAG	CCCCAGCTTGGACATCCCACTCCCTT	CCCAAACTGGGGCCCTCTGAG	---613		
DB	191	---Ala	ProProGlyAsnProAlaGlyPheArg	GluLeu---LeuSerGluValLeu	208		
QY	614	-----A	ACCACTGAAGCGCGCTGTGGTGCCG	---GGGGAAGAG	---649		
DB	209	Glu	ProGlyProLeuProAlaSerLeuProProAla	GlyGluGlnLeuLeuProAspLeu	228		
QY	650	-----TGG	AGTTCGAGTGCAGCGCTTCTACCGGGGC	682			
DB	229	Leu	lSerProHisMetLeuProLeuThrAspLeu	GluIleIysPheGlnTyrArgGly	248		
QY	683	CGCA	AGTCTTCCAGCAGACCATCTCCTCG	CGGAGGCGCTCGCGTGTGGGGTCCGAA	742		
DB	249	Arg	ProArgAlaLeuThrIleSerAsnProHis	GlyCysArgLeuPheTyrSerGln	268		
QY	743	GTG	-----GGAGACAGGACGCTGCCTG	GATGGCCA	772		
DB	269	Leu	GluAlaThrGlnGlnValGluLeuPhe	GlyProIleSerLeuGlu	---Gln	286	
QY	773	GT	CACATCCAGACCTCGCATGCTCC	TGACAGACAGGGAGTAGTACGTACGTGAGG	832		
DB	287	Val	ArgPheProSerProGluAspIleProSer	AspLys---GlnArgPheTyrThrAsn	305		
QY	833	CAT	GTCTGAGTGTCTGGTGGGGACTGGCT	CTCTGCGCGCGCGGAGGCTCTCG	892		
DB	306	Gln	LeuLeuAspValLeuAspArgGlyLeu	IleLeuGlnLeuGlnGlyGlnAspLeuTyr	325		
QY	893	GCC	AGCGCTGGGCACTGCCACATAC	TGG-----GCATGTAGCGAGGAG	940		
DB	326	Ala	IleArgLeuCysGlnCysLysValPheTyr	SerGlyProCysAlaSerAlaHisAsp	345		
QY	941	CTG	CTCCCCAACAGCGGGCATGGCGCT	GTATGCGGAGTCCCCAAGGACAAAGGAGGCG	1000		
DB	346	Ser	CysProAsn-----Pro	IleGlnArgGlyValIlylThrLys	358		
QY	1001	GTG	TTGACCTGGGCGCCTTCATTTAGAT	CTGATTACCTTCCAGGAGGA	---AGCGGA	1057	
DB	359	Leu	PheSerLeuGluHisPheLeuAsnGlu	IleLeuPheGlnIlylGlyGlnThrAsn	378		
QY	1058	CG	TCACACGCTATGCCCTCTGTTCTGT	GTGGGAGTATGCCCCAGACACGCGC	1117		
</							

Fri Oct 29 11:11:45 2004

```

QY 1178 ATGCGCCGGTAGGGGTGCTCTCTCCCTGGAGAACTCTGACCTGCGCATTTTCCAC 1237
Db 1178 ATGCGCCGGTAGGGGTGCTCTCTCCCTGGAGAACTCTGACCTGCGCATTTTCCAC 1237
QY 419 MetPheSer---GlyGluLeuSerTrpSerAlaAspSerIleArgLeuGlnIleSerAsn 437
Db 419 MetPheSer---GlyGluLeuSerTrpSerAlaAspSerIleArgLeuGlnIleSerAsn 437
QY 1238 AGCCA-----CCCACTCTCCCTCAC-----1257
Db 1238 AGCCA-----CCCACTCTCCCTCAC-----1257
QY 438 --ProAspLeuLeuAspArgMetValGluGlnPheLysGluLeuHisIleTrpGlns 457
Db 438 --ProAspLeuLeuAspArgMetValGluGlnPheLysGluLeuHisIleTrpGlns 457
QY 1258 -----CTCGACCACTACAGGCTTACCTGAGGAGCTTGGTGGAGGGCATGGAT 1306
Db 1258 -----CTCGACCACTACAGGCTTACCTGAGGAGCTTGGTGGAGGGCATGGAT 1306
QY 457 erGlnGlnArgLeuGlnProValAlaGlnAlaProGlyAlaGlyLeuGlyValGly- 476
Db 457 erGlnGlnArgLeuGlnProValAlaGlnAlaProGlyAlaGlyLeuGlyValGly- 476
QY 1307 TTCAGGGCCCTGG 1320
Db 1307 TTCAGGGCCCTGG 1320
QY 477 --GlnGlyProTrp 480
Db 477 --GlnGlyProTrp 480

RESULT 11
AAR90326 PRELIMINARY; PRT; 488 AA.
AC AAR90326;
DT 02-MAR-2004 (TREMELrel. 27, Created)
DT 02-MAR-2004 (TREMELrel. 27, Last sequence update)
DT 02-MAR-2004 (TREMELrel. 27, Last annotation update)
DE Interferon regulatory factor 5 variant 4.
GN IRF5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-21316468; PubMed-11303025;
RA Barnes B.J., Moore P.A., Pitha P.M.;
RT "Virus-specific activation of a novel interferon regulatory factor,"
RT IRF-5, results in the induction of distinct interferon alpha genes.";
RL J. Biol. Chem. 276:23382-23390 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Barnes B.J., Pitha P.M.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY504947; AAR90326.1; -.
SQ SEQUENCE 488 AA; 54942 MW; 4E43B54EE74A7B37 CRC64;

Alignment Scores:
Pred. No.: 8,74e-21 Length: 488
Score: 504.50 Matches: 163
Percent Similarity: 44.47% Conservative: 62
Best Local Similarity: 32.21% Mismatches: 186
Query Match: 18.69% Indels: 95
DB: 2 Gaps: 22

US-09-975-253-1 (1-1426) x AAR90326 (1-488)
QY 2 GTTCCAGCTGCGGACGCCCGGACCTTCCATCTGAGCGCGACCATGGGAACCCCAAG 61
Db 2 GTTCCAGCTGCGGACGCCCGGACCTTCCATCTGAGCGCGACCATGGGAACCCCAAG 61
QY 5 IleProValAlaProTrpProArgVal-----15
Db 5 IleProValAlaProTrpProArgVal-----15
QY 62 CCACGNTCTGCGCTGCTGTGCGAGCTGCGACCTGGGCGCACTGGAGGGCGTGCC 121
Db 62 CCACGNTCTGCGCTGCTGTGCGAGCTGCGACCTGGGCGCACTGGAGGGCGTGCC 121
QY 16 ---ArgLeuLysProTrpLeuValAlaGlnValAsnSerCysGlnTyProGlyLeuGln 34
Db 16 ---ArgLeuLysProTrpLeuValAlaGlnValAsnSerCysGlnTyProGlyLeuGln 34
QY 122 TGGGTGAACAAGAGCGCGACCGCTTCCGATCCCTTGGAGACGAGCGGCTAGCGAGGAT 181
Db 122 TGGGTGAACAAGAGCGCGACCGCTTCCGATCCCTTGGAGACGAGCGGCTAGCGAGGAT 181
QY 35 TrpValAsnGlyGluLysLysLeuPheCysIleProTrpArgHisAlaThrHisGly 54
Db 35 TrpValAsnGlyGluLysLysLeuPheCysIleProTrpArgHisAlaThrHisGly 54
QY 182 GCACAGCGAGGAG---GATTTCGGAATCTTCCAGGCTGGCGGAGCGGACCTGGTGCATAT 238
Db 182 GCACAGCGAGGAG---GATTTCGGAATCTTCCAGGCTGGCGGAGCGGACCTGGTGCATAT 238
QY 55 ProSerGlnAspGlyAspAsnThrIlePheLysAlaTrpAlaLysLysLysTyr 74
Db 55 ProSerGlnAspGlyAspAsnThrIlePheLysAlaTrpAlaLysLysLysTyr 74
QY 239 GTTCCCGGAGGAGGATAAGCGACGACCTGCGCAACCTGGAGAGGAATTCGCTCTGCCCTC 298
Db 239 GTTCCCGGAGGAGGATAAGCGACGACCTGCGCAACCTGGAGAGGAATTCGCTCTGCCCTC 298
QY 75 ThrGluGlyValAspGluAlaAspProAlaLysTrpLysAlaAsnLeuArgCysAlaLeu 94
Db 75 ThrGluGlyValAspGluAlaAspProAlaLysTrpLysAlaAsnLeuArgCysAlaLeu 94

```

```

QY 299 AACCGCAAAAGAGGGTGGCTTTAGCAGAGGACCGGAGCAGGAC---CCTCAGACCCA 355
Db 299 AACCGCAAAAGAGGGTGGCTTTAGCAGAGGACCGGAGCAGGAC---CCTCAGACCCA 355
QY 95 AsnLysSerArgAspPheArgLeuIleTyArgGlyProArgAspMetProGlnPro 114
Db 95 AsnLysSerArgAspPheArgLeuIleTyArgGlyProArgAspMetProGlnPro 114
QY 356 CATAAATCTACAGTTTGTGAACCTCAGGAGTTGGGACCTTTCCAGCAGCAGACCTCT 415
Db 356 CATAAATCTACAGTTTGTGAACCTCAGGAGTTGGGACCTTTCCAGCAGCAGACCTCT 415
QY 115 TyrlsileTyrlGluValCysSerAsnGlyPro-----AlaProThrAspSerGln 131
Db 115 TyrlsileTyrlGluValCysSerAsnGlyPro-----AlaProThrAspSerGln 131
QY 416 CCG-----CACCAATGTGTGGAGGACGAGTCTCTGATACCCAGGAGACATCTTGAT 469
Db 416 CCG-----CACCAATGTGTGGAGGACGAGTCTCTGATACCCAGGAGACATCTTGAT 469
QY 132 ProProGluAspTyrlSerPheGlyAlaGlyGluGluGluGluGluGluGln 151
Db 132 ProProGluAspTyrlSerPheGlyAlaGlyGluGluGluGluGluGluGln 151
QY 470 GAGTTACTGGGTAACATG-----GTCTTGGCCCCA 499
Db 470 GAGTTACTGGGTAACATG-----GTCTTGGCCCCA 499
QY 152 ArgMetLeuProSerLeuSerLeuThrGluAspValIlystrProProThrLeuGlnPro 171
Db 152 ArgMetLeuProSerLeuSerLeuThrGluAspValIlystrProProThrLeuGlnPro 171
QY 500 CTCCAGATCCGGACCCCAAGCTGCTGTAGCCCTGAGCCCTGAGCCCTGCTCAGCCCTG 559
Db 500 CTCCAGATCCGGACCCCAAGCTGCTGTAGCCCTGAGCCCTGAGCCCTGCTCAGCCCTG 559
QY 172 --ProThrLeuGlnProProValValLeuGlyProProAlaProAspProSerProLeu 190
Db 172 --ProThrLeuGlnProProValValLeuGlyProProAlaProAspProSerProLeu 190
QY 560 CGGAGCCCGAGCTTGGACATCCCACTCCCTTCCAAACCTGGGGCCCTCTGAG----- 613
Db 560 CGGAGCCCGAGCTTGGACATCCCACTCCCTTCCAAACCTGGGGCCCTCTGAG----- 613
QY 191 ---AlaProProProGlyAsnProAlaGlyPheArgGluLeu---LeuSerGluValLeu 208
Db 191 ---AlaProProProGlyAsnProAlaGlyPheArgGluLeu---LeuSerGluValLeu 208
QY 614 -----AAACCCACTGAAGCGGCTGTTGTCGG---GGGGAAGAG----- 649
Db 614 -----AAACCCACTGAAGCGGCTGTTGTCGG---GGGGAAGAG----- 649
QY 209 GluProGlyProLeuProAlaSerLeuProAlaGlyGluGlnLeuLeuProAspLeu 228
Db 209 GluProGlyProLeuProAlaSerLeuProAlaGlyGluGlnLeuLeuProAspLeu 228
QY 650 -----TGGGAGTTTCGAGGTGACAGCTTCTTACCGGGGC 682
Db 650 -----TGGGAGTTTCGAGGTGACAGCTTCTTACCGGGGC 682
QY 229 LeuIleSerProHisMetLeuProLeuThrAspLeuGluIleLysPheGlnTyArgGly 248
Db 229 LeuIleSerProHisMetLeuProLeuThrAspLeuGluIleLysPheGlnTyArgGly 248
QY 683 CGCCAAAGTCTTCCAGCAGACCATCTCTCCCGAGGGGCTGCGCTGTGGGTCGAA 742
Db 683 CGCCAAAGTCTTCCAGCAGACCATCTCTCCCGAGGGGCTGCGCTGTGGGTCGAA 742
QY 249 ArgProProArgAlaLeuThrIleSerAsnProHisGlyCysArgLeuPheTyrlSerGln 268
Db 249 ArgProProArgAlaLeuThrIleSerAsnProHisGlyCysArgLeuPheTyrlSerGln 268
QY 743 GTG-----GGAGACAGACGCTGCTGGATGGCCCA 772
Db 743 GTG-----GGAGACAGACGCTGCTGGATGGCCCA 772
QY 269 LeuGluAlaThrGlnGluGlnValGluLeuPheGlyProIleSerLeuGlu-----Gln 286
Db 269 LeuGluAlaThrGlnGluGlnValGluLeuPheGlyProIleSerLeuGlu-----Gln 286
QY 773 GTACACTGCCAGACCTTGGCATGTCCCTGACACAGACAGGGGAGTGTAGTACGTGAGG 832
Db 773 GTACACTGCCAGACCTTGGCATGTCCCTGACACAGACAGGGGAGTGTAGTACGTGAGG 832
QY 287 ValArgPheProSerProGluAspIleProSerAspLys---GlnArgPheTyrlThrAsn 305
Db 287 ValArgPheProSerProGluAspIleProSerAspLys---GlnArgPheTyrlThrAsn 305
QY 833 CATGTGCTGAGCTGCTGCTGGTGGGAGCTGCTCTCTGGCGGGCGCGGAGTGGCTCTGG 892
Db 833 CATGTGCTGAGCTGCTGCTGGTGGGAGCTGCTCTCTGGCGGGCGCGGAGTGGCTCTGG 892
QY 306 GlnLeuLeuAspValLeuAspArgGlyLeuIleLeuGlnLeuGlnAspLeuTyrl 325
Db 306 GlnLeuLeuAspValLeuAspArgGlyLeuIleLeuGlnLeuGlnAspLeuTyrl 325
QY 893 GCCCAGCGGCTGGGCGACTGCCACACATCTGG-----GCAGTGAGCGGAGGAG 940
Db 893 GCCCAGCGGCTGGGCGACTGCCACACATCTGG-----GCAGTGAGCGGAGGAG 940
QY 326 AlaIleArgLeuCysGlnCysLysValPheTrpSerGlyProCysAlaSerAlaHisAsp 345
Db 326 AlaIleArgLeuCysGlnCysLysValPheTrpSerGlyProCysAlaSerAlaHisAsp 345
QY 941 CTGCTCCCAACAGCGGGCATGGGCGCTGATGGCGAGGTCCCAAGGACAGGAGGAGGC 1000
Db 941 CTGCTCCCAACAGCGGGCATGGGCGCTGATGGCGAGGTCCCAAGGACAGGAGGAGGC 1000
QY 346 SerCysProAsn-----ProIleGlnArgGluValIlystrLys 358
Db 346 SerCysProAsn-----ProIleGlnArgGluValIlystrLys 358
QY 1001 GTGTTTTCACCTGGGCGCCCTTCAATTGTAGATCTGTATCTTACCTTACCGGAGGA---AGCGGA 1057
Db 1001 GTGTTTTCACCTGGGCGCCCTTCAATTGTAGATCTGTATCTTACCTTACCGGAGGA---AGCGGA 1057
QY 359 LeuPheSerLeuGluHisPheLeuAsnGluLeuIleLeuPheGlnLysGlyGlnThrAsn 378
Db 359 LeuPheSerLeuGluHisPheLeuAsnGluLeuIleLeuPheGlnLysGlyGlnThrAsn 378
QY 1058 CGCTCACACGCTATGCGCTCTGCTTGTGTGGGGAGTGCATGGCCCGCAGGACGAGCG 1117
Db 1058 CGCTCACACGCTATGCGCTCTGCTTGTGTGGGGAGTGCATGGCCCGCAGGACGAGCG 1117
QY 379 ThrProProProGluIlePhePheCysPheGlyGluGluTrpProAspArgLysPro 398
Db 379 ThrProProProGluIlePhePheCysPheGlyGluGluTrpProAspArgLysPro 398
QY 1118 TGGACCAAGAGCTGCTGATGCTCAAGGTTCTGCGCCACGCTGCTCAGGCGCTTGTGTA 1177
Db 1118 TGGACCAAGAGCTGCTGATGCTCAAGGTTCTGCGCCACGCTGCTCAGGCGCTTGTGTA 1177
QY 399 ArgGluLysLysLeuIleThrValGlnValProValAlaAlaAlaLeuLeuGlu 418
Db 399 ArgGluLysLysLeuIleThrValGlnValProValAlaAlaAlaLeuLeuGlu 418
QY 1178 ATGGCCCGGGTAGGGGTGCGCTCTCTCTGGAGAACTACTGTGACCTGACATTTCCAC 1237
Db 1178 ATGGCCCGGGTAGGGGTGCGCTCTCTCTGGAGAACTACTGTGACCTGACATTTCCAC 1237
QY 419 MetPheSer---GlyGluLeuSerTrpSerAlaAspSerIleArgLeuGlnIleSerAsn 437
Db 419 MetPheSer---GlyGluLeuSerTrpSerAlaAspSerIleArgLeuGlnIleSerAsn 437
QY 1238 AGCCA-----CCCACTCTCCCTCAC-----1257
Db 1238 AGCCA-----CCCACTCTCCCTCAC-----1257

```


Db	27	SerCysGlnTyrProGlyLeuGlnTrpValAsnGlyGluLeuLysLeuPheTyrIlePro	46
QY	158	TGAAACACAGCGCTACGGCAGGATGCACAGCAGGAG--GATTTCCGATCTTCCAGGCC	214
Db	47	TtpArgHisAlaThrArgHisGlyProSerGlnAspGlyAspAsnThrIlePheLysAla	66
QY	215	TGGCCCGAGGCCACTGGTGCATATGTTCCCGGAGGGATAGCCAGACCTGCCAACCTGG	274
Db	67	TtpAlaLysGluThrGlyLysTyrThrGluGlyValAspGluAlaAspProAlaLysTrp	86
QY	275	AAAGAGATTTCCGCTCTGCCCTCAACCGCAAGAGGGTTGCGTTAGCAGACGACCGG	334
Db	87	LysAlaAsnLeuArgCysAlaLeuAsnLysSerArgAspPheGlnLeuPheTyrAspGly	106
QY	335	AGCAAGGAC--CCTCAGCACCACATAAAATCTACGAGTTGTGAACTCAGGATTTGGG	391
Db	107	ProArgAspMetProGlnProTyrIleTyrGluValCysSerAsnGlyProAla	126
QY	392	GACTTT--TCCAGCCAGACACCTCTCCGACCAACCAATGCTGGAGGAGTACTTCTCAT	448
Db	127	ProThrGluSerGlnPro-----ThrAspAspTyrValLeuGlyGluGluGluGln	143
QY	449	ACCAGGAAGACATTCTGGATGAGTTACTGGGTAAACATGTTGTGGTCCCACTCCACAT	508
Db	144	GluGluGluGluGluGlnLeuGlnArgMetLeuProGlyLeuSerIleThrGluProAlaLeu	163
QY	509	CGGGACCCCA-----AGCTGGCTGTAGCCCTGAGCCCTGCCCTCAGCCCC	556
Db	164	ProGlyProProAsnAlaProTyrSerLeuProLysGluAspThrLysTrpProAla	183
QY	557	CTGGGAGCCCC--AGCTTGGCAATCCCACTCCCTCCCAACCTG-----GGGCCC	607
Db	184	LeuGlnProProValGlyLeuGlyProProValProAspProAsnLeuAlaProPro	203
QY	608	TCTGAGAACCA-----CTGAAGCGGCTGTGGTGGCGGG-----643	643
Db	204	SerGlyAsnProAlaGlyPheArgGlnLeuLeuProGluValLeuGluProGlyProLeu	223
QY	643	-----643	643
Db	224	AlaSerSerGlnProProThrGluProLeuLeuProAspLeuLeuIleSerProHisMet	243
QY	644	--GAAGAGTGGAGTTCGAGGTGCAGCCTTCTACGGGGCGGCCAAGTCTTCCAGCAG	700
Db	244	LeuProLeuThrAspLeuGluIleLysPheGlnTyrArgGlyArgAlaProArgThrLeu	263
QY	701	ACCATCTCTGCCGAGAGGCTCGCGCTGTGGGTCCGAGTG-----745	745
Db	264	ThrIleSerAsnProGlnGlyCysArgLeuPheTyrSerGlnLeuGluAlaThrGlnGlu	283
QY	746	-----GGAGACAGACGCTGCCTGGATGCCAGTCCACATGCACAGACCTC	790
Db	284	GlnValGluLeuPheGlyProValThrLeuGlu-----GlnValArgPheProSerPro	301
QY	791	GGCATGTCCCTGCACACAGGGGAGTGATGACTAGCTAGGCGCATGTGCTGAGCTGCCTG	850
Db	302	GluAspIleProSerAspLys---GlnArgPheTyrThrAsnGlnLeuLeuAspValLeu	320
QY	851	GTTGGGGGACTGGCTCTCTGCGGGCGGGCAGTGGCTCTGGCCACCGCTGGGGCAC	910
Db	321	AspArgGlyLeuIleLeuGlnLeuGlnGlyGlnAspLeuTyrAlaIleArgLeuCysGln	340
QY	911	TGCCACACATCTGG-----GCAGTGGACGAGGAGTCTGCTCCCAACACGCGG	958
Db	341	CysLysValPheTrpSerGlyProCysAlaLeuAlaHisGlySerCysProAsn-----	358
QY	959	CATGGGCTGTATGGGAGGTCCCAAGGACAAAGAGGAGGCGGTGTTGACTCTGGGGCCC	1018
Db	359	-----ProIleGlnArgGluValLysThrLysLeuPheSerLeuGlnGln	373
QY	1019	TTCATTGTAGATCTGATTACCTTCACGGAAGA---ACGGAGCGTCCACAGCTATGCC	1075

QY	557	CTGCGGAGCCCC	---	AGCTTGGA	CAATCCCACTCCCTCTCCCAACCTG	-----GGGCCC	607
Db	184	LeuGlnProValGlyLeuGlyProValProAspProAsnLeuLeuAlaProPro	203				
QY	608	TCTGAGAACCCA	-----CTGAAGCGGTGTGGTGGCGGG	-----	643		
Db	204	SerGlyAsnProAlaGlyPheArgGlnLeuLeuProGluValLeuGluProGlyProLeu	223				
QY	643	-----	-----	-----	-----	-----	643
Db	224	AlaSerSerGlnProProThrGluProLeuLeuProAspLeuLeuIleSerProHisMet	243				
QY	644	---GAAGAGTGGGAGTTCGAGGTGCAGCGCTTCTACCGGGCGCGCAAGTCITCCACGAG	700				
Db	244	LeuProLeuThrAspLeuGluIleLysPheGlnTyrArgLysArgAlaProArgThrLeu	263				
QY	701	ACCATCTCTCGCGGAGCGCTGCGCTGTGGGTCCGAGTG	745				
Db	264	ThrIleSerAsnProGlnGlyCysArgLeuPheTyrSerGlnLeuAlaThrGlnGlu	283				
QY	746	-----GGAGACAGACGCTGCCTGGATGCCAGTCACACTGCCAGACCCCT	790				
Db	284	GlnValGluLeuPheGlyProValThrLeuGlu	301				
QY	791	GGCATCTCCCTGACACACAGGGAGTGATGACCTAGTGAGCATGTCTGAGCTGCCTG	850				
Db	302	GluAspIleProSerAspLys	320				
QY	851	GGTGGGGAGTGGCTCTCTGGCGGGCGGGACAGTGGCTCTGGGCCACAGCGGTGGGGCAC	910				
Db	321	AspArgGlyLeuIleLeuGlnGlnGlyGlnAspLeuTyrAlaIleArgLeuCysGln	340				
QY	911	TGCCACACATACTGG	958				
Db	341	CysLysValPheTrpSerGlyProCysAlaLeuAlaHisGlySerCysProAsn	358				
QY	959	CATGGCCCTGATGGCGAGTCCCCAAGACACAGGAAGGAGCGTGTTCGACTGGGGCCC	1018				
Db	359	-----ProileGlnArgGluValLysThrLysLeuPheSerLeuGlnGln	373				
QY	1019	TTCATTGTAGATCTGATTACCTTCACGGAAGA	1075				

Qy	643	-----	643
Db	224	AlaSerSerGlnProProThrGluProLeuLeuProAspLeuLeuLeuSerProHisMet	243
Qy	644	---CAACAGCTGGGAGTTCGAGGTACAGCCTTCTACCGGGGCGCCCAAGTCTTCCAGCAG	700
Db	244	LeuProLeuThrAspLeuGluIleLysPheGlnTyrArgGlyArgAlaProArgThrLeu	263
Qy	701	ACCATTCTCTCGCCGAGGAGCCCTCGCGCTGGTGGGTCCGAAGTG-----	745
Db	264	ThrIleSerAsnProGlnGlyCysArgLeuPheTyrSerGlnLeuGluAlaThrGlnGlu	283
Qy	746	-----GGACAGCAGCAGCTGCCTGGATGCCACGTACACTGCCAGACCCCT	790
Db	284	GlnValGluLeuPheGlyProValThrLeuGlu-----GlnValArgPheProSerPro	301
Qy	791	GGCATCTCCCTGACACACAGGGGAGTGATGACTACGTACGATCGTCTGCTGAGCTGCCTG	850
Db	302	GluAspIleProSerAspLys---GlnArgPheTyrThrAsnGlnLeuLeuAspValLeu	320
Qy	851	GGTGGGGAGTGGCTCTTGGCGGGCCGGCAGTGGCTCTGGGCCCGCAGCGCTGGGGCAC	910
Db	321	AspArgGlyLeuIleLeuGlnGlnGlyGlnAspLeuTyrAlaIleArgLeuCysGln	340
Qy	911	TGCCACACATACTGG-----GCAGTGCAGCGAGGAGCTCTCCCAACACGCGG	958
Db	341	CysLysValPheTrpSerGlyProCysAlaLeuAlaHisgIysSerCysProAsn-----	358
Qy	959	CATGGGCGCTGATGGCGAGTCCCAAGACACAGGAAGGAGGCGTGTTCAGCTGGGGCCC	1018
Db	359	-----ProIleGlnArgGluValLysThrLysLeuPheSerLeuGlnGln	373
Qy	1019	TTCATTGTAGATCTCATCTTACCGGAAGA---ACGGCAGCGTCACACAGCTATGCC	1075

Db	302	Gl uAspIleProSerAspLys---GlnArgPheTyrThrAsnGlnLeuLeuAspValIleu	320
QY	851	GGTGGGGGACTGGCTCTCTGGCGGGCCGGAGTGTCTGGGCGCCAGCGGTGGGGCAC	910
Db	321	AspArgGlyLeuIleuLeuGlnGlnGlyGlnAspLeuTyrAlaIleArgLeuCysGln	340
QY	911	TGCCACACATACTGG-----GCAGTGACGCGAGGAGCTCTCCCCACACAGCGGG	958
Db	341	CysLysValPheTrpSerClyProCysAlaIleuAlaHisGlySerCysProAsn-----	358
QY	959	CATGGGCGTGATGGCGAGTCCCAAGACACAGGAAGGAGGCGTGTTTGACTCGGGGCC	1018
Db	359	-----ProIlelnArgIluValIysThrLysLeuPheSerLeuGluGln	373
QY	1019	TTCAATTGATGTCATTACCTTCACGGAAGA--ACGGGACGCTCACCGCTATGCC	1075

[illegible]

374	Db	PheLeuAsnGluLeuLeuPheGlnLysGlyGlnThrAsnThrProProPheGlu	393
1076	Qy	CTCTGGTTCTGTGGGGAGTCA TGCCCCAGAGACAGCCGTGGACCAAGAGCTCGTG	1135
394	Db	IlePhePheCysPheGlyGluGluTrpProAspValLysProArgGluLysLysLeuIle	413
1136	Qy	ATGTCACAGTTGTGCCACGTGCTCAGCGCCTTGTTAGAAATGGCCCGGTAGCGGGT	1195
414	Db	ThrValGlnValValProValAlaAlaArgLeuLeuLeuGluMetPheSer--GlyGlu	432
1196	Qy	GCCTCTCCCTGGAGAACTACTGTGGACCTGCACATTTCCACAGCACCCACTCTCCCTC	1255
433	Db	LeuSerTrpSerAlaAspSerIleArgLeuGlnIleSerAsnProAspLeuLysAspHis	452
1256	Qy	ACCTCCGACCACTAGTCAAG-----GCTACCTCCAGGAC	1288
453	Db	MetValGluGlnPheLysGluLeuHisIleLeuTrpGlnSerGlnGlnLeuGlnPro	472
1289	Qy	TTGGTGGAG-----GGCATGGAT--TTCCAGGGCCCT	1318
473	Db	MetValGlnAlaProValAlaGlyLeuAspAlaSerGlnGlyPro	498

RESULT 13

IRF5	HUMAN	STANDARD;	PRT;	498 AA.
ID	IRF5_HUMAN			
AC	Q13568; Q9BQF0;			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	05-JUL-2004 (Rel. 44, Last annotation update)			
DE	Interferon regulatory factor 5 (IRF-5).			
GN	Name=IRF5;			
OS	homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
NCBI_TaxID=9606;				
[1]	SEQUENCE FROM N.A.			
RP	Grossman A., Mitrucker H.W., Lantonio L., Mak T.W.;			
RA	Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.			
RL	[2]			
RN	SEQUENCE FROM N.A.			
RP				

TISSUE=Kidney; PubMed=12477932; DOI=10.1073/pnas.242603899;

RC MEDLINE=22388257; Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA RAUSBERG R.L., Feingold E.A., Grouse L.H., Schenmen C.M., Schuler G.D.,
RA Klausner R.D., Collins F.S., Wagner L., Schenmen C.F., Bhat N.K.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Hsieh F.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hong L.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hsing L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J.J., Abramson R.D., Mullihay S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huliyil S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Keckman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grinstead J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerrach A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences." Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SUBCELLULAR LOCATION: Nuclear.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

```

-----
CC      EMBL; U51127; AAA96056.1; -.
DR      EMBL; BC004201; AAH04201.1; -.
DR      EMBL; BC004139; AAH04139.1; -.
DR      PIR; G02474; G02474.
DR      HSSP; P23906; 1IRF.
DR      TRANSFAC; T05104; -.
DR      Genew; HGNC:6120; IRF5.
DR      MIM; 607218; -.
DR      InterPro; IPR001346; IRF.
DR      InterPro; IPR008984; SMAD_FHA.
DR      InterPro; IPR009058; Wing_hlx_DNA_bnd.
DR      Pfam; PFO0605; IRF; 1.
DR      PRINTS; PR00267; INTERNEGCT.
DR      ProDom; PD002355; IRF; 1.
DR      SMART; SM00348; IRF; 1.
DR      PROSITE; PS00601; IRF; 1.
DR      DNA-binding; Nuclear protein; Transcription regulation.
KW      DNA BINDING; 16 118      tryptophan pentad repeat.
FT      DOMAIN      142 149      poly-Glu.
FT      CONFLICT    161 175      EDVWMPITQPPTLR -> DAVQSGPHNTFYSLLKEDVW
FT                                     (in Ref. 1).
SQ      SEQUENCE 498 AA; 56044 MW; 01B2ED95C28384E8 CRC64;

Alignment Scores:
Pred. No.: 1.69e-20
Score: 499.50
Percent Similarity: 43.60%
Best local Similarity: 31.59%
Query Match: 18.51%
Gaps: 1
DB: 1-498

```

UIC 00 075-353-1 (1-1426) X IRF5 HUMAN (1-498)

2	GTTCAGACTCCCGCAGCGCCGACCTTCATCGTAGCCGGACCATGGGAACCCCAAAG	61
Qy		
5	IleProValAlaProThrProProArgArgVal	15
Db		
62	CCAGGNTCTCGCTGGCTGGTGTGCAGCTGGACCTGGGGCACTGGAGGCGGTGGCC	121
Qy		
16	---ArgLeuIysProTrpIleuValAlaGlnValAsnSerCysGlnTyrProGlyLeuGln	34
Db		
122	TGGGTGAACAAGAGCCGACGCGCTCCGCATCCCTTGGAGACACGCCCTACGCACGAT	181
Qy		
35	TrpValAsnGlyGluIysLysLeuPheCysIleProTrpArgHisAlaThrArgHisGly	54
Db		
182	GCACAGCAGGAG---GATTCGGAACTCTTCAGCCCTGGGCGGAGGCACCTGGTGCATAT	238
Qy		
55	ProSerGlnAspGlyAspAsnThrIlePheIysAlaTrpAlaIysGluThrGlyIysTyr	74
Db		
239	GTTCCCGGAGGAGTAAGCCAGACCTGCCAACCTTGGAGAGGAATTCCTCGCTCTGCCTC	298
Qy		
75	ThrGluGlyValAspGluIysAlaAspProAlaIysTrpIysAlaAsnLeuArgCysAlaLeu	94
Db		
299	AACCCCAAGAAGGTTTCGGTTTAGCAGAGGACCGGACCAAGGAC---CCTCAGCACCCA	355
Qy		
95	AsnLysSerArgAspPheArgIleuIleTyrAspGlyProArgAspMetProGlnPro	114
Db		
356	CATAAAATCTACGATTTGTGAACCTCAGAGTGTGGGACTTTTCCACGCGACACCTCT	415
Qy		
115	TyrLysIleTyrGluValCysSerAsnGlyPro-----AlaProThrAspSerGln	131
Db		
416	CGG-----GACACCAATGGTCGAGGAGTACTCTGTATACCACGAGACACATTCTGGAT	469
Qy		
132	ProProGluAspTyr-SerPheGlyAlaGlyGluGluGluGluGluGluGluGln	151
Db		
470	GAGTACTGGGTAAACATG-----	487
Qy		
152	ArgMetLeuProSerLeuSerLeuThrGluAspValYsTrpProProThrLeuGlnPro	171
Db		
488	-----GTGTTGGCCCACTCCAGATCCCGGACCCCGACCTGGCT	529
Qy		
172	ProThrLeuArgProProThrLeuGlnPro---ProThrLeuGlnProProValValLeu	190
Db		

QY	530	GTAGCCCTGAGCCCTGCCTCAGCCCTCGGAGCCCGAGCTTGAGCAATCCCACTCCC	589
Db	191	GlyProProAlaProAspSerProLeu---AlaProProProGlyAsnProAlaGly	209
QY	590	TTCCCAACCTGGGGCCCTCTGAG---	637
Db	210	PheArgGluLeu---LeuSerGluValLeuGluProGlyProLeuProAlaSerLeuPro	228
QY	638	CCG---GGGGAAGAG---	652
Db	229	ProAlaGlyGluGlnLeuLeuProAspLeuLeuLeuSerProHisMetLeuProLeuThr	248
QY	653	GAGTTCGAGGTGCACAGCCCTTCTACCGGGCGCCCAAGTCTTCCAGCAGACCATCTCCCTGC	712
Db	249	AspLeuGlnIleLysPheGlnTyArgGlyArgProProArgAlaLeuThrIleSerAsn	268
QY	713	CCGAGGGCCTGGCTGGTGGGTCCGAAGTG---	745
Db	269	ProHisGlyCysArgLeuPheTySerGlnLeuGluAlaThrGlnGluGlnValGluLeu	288
QY	746	---GGAGCAGACGCTGCCTGGATGGCAGTCCACTGCCAGACCCCTGGCATGTCCCTG	802
Db	289	PheGlyProIleSerLeuGlu-----GlnValArgPheProSerProGluAspIlePro	306
QY	803	ACAGACAGGGAGTGTAGCTACGTGAGGCATGTGTGACTGCCTGGTGGGGGAGCTG	862
Db	307	SerAspLys---GlnArgPheTyThrAsnGlnLeuLeuAspValLeuAspArgGlyLeu	325
QY	863	GCTCTCTGGGGGCCGGGCAGTGGCTCTGGGCCAGCGGTGGGGGACTGCCACACATCA	922
Db	326	IleLeuGlnLeuGlnGlyGlnAspLeuTyraIaIleArgLeuCysGlnCysLysValPhe	345
QY	923	TGG-----GCAGTGCAGCGAGGAGTGCTCCCCAAACGCGGCATGGCCCTGAT	970
Db	346	TrpSerGlyProCysAlaSerAlaHisAspSerCysProAsn-----	359
QY	971	GGCGAGTCCCCAAGACAAGAGGAGGCGGTGTTGACCTGGGGCCCTTCATTGTAGAT	1030
Db	360	---ProIleGlnArgGluValLysThrLysLeuPheSerLeuGluHisPheLeuAsnGlu	378
QY	1031	CTGATTACCTTTCACGGAAGCA---AGCGGAGCGTCCACAGCTATGCCCTCTGGTTCTGT	1087
Db	379	LeuIleLeuPheGlnLysGlyGlnThrAsnThrProProPheGluIlePhePheCys	398
QY	1088	GTGGGGAGTCTAGCCCCCAGGACCCGCTGGACCAAGAGGCTCGTGTGTCGTCACAGGT	1147
Db	399	PheGlyGluGluTrpProAspArgLysProArgGluLysLysLeuIleThrValGlnVal	418
QY	1148	GTCCCCAGCTGCTCAGGGCCTTGTAAGAATGGCCCGGTAGGGGTGCTCTCCCTG	1207
Db	419	ValProValAlaAlaArgLeuLeuLeuLeuMetPheSer---GlyGluLeuSerTrpSer	437
QY	1208	GAGAATACTGTGACCTGCACATTCACACAGCCA-----	1242
Db	438	AlaAspSerIleArgLeuGlnIleSerAsn---ProAspLeuLysAspArgMetValGluG	457
QY	1243	-----CCCCTCTCCCTCAC-----	1276
Db	457	lnPheLysGluLeuHisIleTrpGlnSerGlnArgLeuGlnProValAlaGlnA	477
QY	1277	TACCTCGAGCATTTGGTGGAGGCATGGATTTCCAGGGCCCTGG	1320
Db	477	laProProGlyAlaGlyLeuGlyValGly---GlnGlyProTrp	490

RESULT 14

Q6PGZ7 PRELIMINARY; PRT; 492 AA.

AC Q6PGZ7; ID

DT 05-JUL-2004 (TrEMBLrel. 27, Created)

DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)

DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

DE Hypothetical protein zgc:63500.

Db 426 SerGluMetPheSerGlyAspCysThrArgSerPheAspSerGlySerValArgLeuGln 445
 QY 1229 ATTTCACAGCAGCACCACTCTCCCTACCTCGACACAGTACAGGCTACCTGCAGGAC 1288
 Db 446 lIeSer-----lIleProAspIleLysAsnIleValThrHisLeuLysGln 461
 QY 1289 TTGGTGGAGGCATGGAT---TTCCAGGGCCCTGGGAGAGCTGAGCCCTCGCTCT 1342
 Db 462 LeuTyArgLeuLeuGlnThrHisGlnGlyPro---GluAsnTrpProLeuProPro 479
 RESULT 15
 AAH56772 PRELIMINARY; PRT; 492 AA.
 ID AAH56772
 AC AAH56772; (TrEMBLrel. 27, Created)
 DT 24-MAY-2004 (TrEMBLrel. 27, Last sequence update)
 DT 24-MAY-2004 (TrEMBLrel. 27, Last annotation update)
 DT DE Hypothetical protein zgc:63500.
 OS ZGC:63500.
 GN Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 NCBI_TaxID:7955;
 RN [1]
 RX SEQUENCE FROM N.A.
 RC STRAIN=AB; TISSUE=Whole body;
 RE MEDLINE=22389257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L.H., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Gay L.J., Hulyk S.W.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Lu X., Gibbs R.A.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Rodriguez S., Sanchez A.,
 RA Fahy J., Hellon E., Kettelman M., Madan A., Rodriguez Y., Bouffard G.G.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smalios D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RL and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RX SEQUENCE FROM N.A.
 RC STRAIN=AB; TISSUE=Whole body;
 RE Strausberg R.;
 RA Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
 RL ENBL; BC056772; AAH56772.1; -.
 DR DR Hypothetical protein.
 KW Hypothetical protein.
 SQ SEQUENCE 492 AA; 54876 MW; 98ER8FD3FE1F589B CRC64;
 Alignment Scores:
 Pred. No.: 2,86e-18 Length: 492
 Score: 460.50 Matches: 146
 Percent Similarity: 42.69% Conservative: 67
 Best Local Similarity: 29.26% Mismatches: 185
 Query Match: 17.06% Indels: 101
 DB: 2 Gaps: 21
 US-09-975-253-1 (1-1426) x AAH56772 (1-492)
 QY 65 CGGNTCTCGCCCTGGCTGGTGTTCAGCTGCACCTGGGCACTGGAGGGCTGGCGCTGG 12
 Db 9 ArgLeuLysProTrpLeuValSerGlnValAspAsnAlaThrPheProGlyLeuValTrp 28
 QY 125 GTGAACAGAGCCGACGCGCTTCGCGATCCCTTGGAGACGCGGCTACGGCAGGATGCA 18

This Page Blank (uspto)